Spectra of Normalized Laplace Operators for Graphs and Hypergraphs

Der Fakultät für Mathematik und Informatik
der Universität Leipzig
eingereichte

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

DOCTOR RERUM NATURALIUM
(Dr. rer. nat.)

im Fachgebiet

Mathematik

vorgelegt

von Raffaella Mulas
geboren am 09.07.1992 in Cagliari, Italien

Betreuer: Prof. Dr. Jürgen Jost

Datum der Einreichung: 11.12.2019

Die Annahme der Dissertation haben empfohlen:
Prof. Dr. Jürgen Jost (MPI MiS Leipzig)
Prof. Dr. Norbert Peyerimhoff (Durham University, UK)

Die Verleihung des akademischen Grades erfolgt mit Bestehen der Verteidigung am 19.06.2020 mit dem Gesamtprädikat *magna cum laude*.

Bibliographische Daten

Spectra of Normalized Laplace Operators for Graphs and Hypergraphs
(Spektren von normalisierten Laplace-Operatoren für Graphen und Hypergraphen)
Mulas, Raffaella
Universität Leipzig, Dissertation, 2020

103 Seiten, 14 Abbildungen, 61 Referenzen

Un grafo è fatto da punti *sparati nello spazio* e connessi tra loro.[1]

(My dad, to whom this thesis is dedicated)

[1] A graph is made by points *shot into space* and connected to each other.

Acknowledgements

I am grateful to Jürgen Jost for being a great source of inspiration, both as a professor and as a person. I thank him for showing me the deep beauty of many branches of mathematics, for the guidance and the support. I thank him for giving me marvellous opportunities, including the one of creating a beautiful network of researchers all around the globe. I am so honored to have as my PhD supervisor such an outstanding mathematician and wonderful person.
I am grateful to Bernd Sturmfels and Ngoc Mai Tran for the inspiration and the support they always give me. I thank them for believing so much in me and for always pushing me to give the best. I thank them for the knowledge and the love for mathematics that they constantly transmit me.
I am grateful to Antonio Lerario and Florentin Münch for giving me the opportunity to collaborate on very interesting projects with wide enthusiasm.
I am grateful to Antje Vandenberg for helping me in every possible situation (even when I cut my finger instead of the onions). I thank her for the laughs, for her generosity, for her kindness and for her incredible energy.
I am grateful to all the PhD students, Postdocs and Professors, both from inside and outside the institute, with whom I had very interesting discussions during these years. In particular I would like to thank (in alphabetical order): Aida Abiad, Pau Aceituno, Eleonora Andreotti, Renan Assimos, Reinhard Diestel, Felix Gaisbauer, Wilmer Leal, Paolo Perrone, Guillermo Restrepo, Sharwin Rezagholi, Emil Saucan.
I am grateful to the Max Planck Institute for Mathematics in the Sciences and to the International Max Planck Research School for the financial support.
I am grateful to Jacinta Torres for bringing magic and beauty at the Institute with her yoga course that ended with these words: *Yoga is like mathematics and like many other things in life. In the beginning you have a teacher who leads you, but at some point you have to start working on your own.*
I am grateful to all the Swing and Blues dancers that during these years have shared either one or hundreds of dances with me. The energy and the beauty that I find on the dance floor has a huge impact on every part of my life, including my work.
Last but not least, I am grateful to my friends and to my beautiful, extended family, for all the love and the support.

Raffaella Mulas

Abstract

This thesis is based on four papers [37, 38, 39, 45] that investigate the spectra of normalized Laplace operators for graphs and hypergraphs.

Fix a graph $\Gamma = (V, E)$ on n vertices that is unweighted, undirected and without loops, multiple edges and isolated vertices. Let Id be the $n \times n$ identity matrix, let A be the *adjacency matrix* of Γ, let D be the diagonal *degree matrix* and let

$$L := \mathrm{Id} - D^{-1}A$$

be the *(normalized) Laplace operator* of Γ. It is known that L has n real, non-negative eigenvalues, counted with multiplicity. We arrange them as

$$\lambda_1 \leq \ldots \leq \lambda_n.$$

Summarizing some classical results on the spectrum of the Laplacian [16],

- $\lambda_1 = 0$ and the multiplicity of the eigenvalue 0 equals the number of connected components of Γ.
- The largest eigenvalue is such that

  $$\lambda_n \geq \frac{n}{n-1}$$

  with equality if and only if Γ is complete, and

  $$\lambda_n \leq 2$$

  with equality if and only if a connected component of Γ is bipartite.
- For connected graphs, the first non-zero eigenvalue λ_2 is controlled both above and below by the *Cheeger constant*, a quantity that measures how difficult it is to partition the vertex set into two disjoint sets V_1 and V_2 such that the number of edges between V_1 and V_2 is as small as possible and such that the *volume* of both V_1 and V_2, i.e. the sum of the degrees of their vertices, is as big as possible.

Summarizing the main results presented in this thesis,

(a) We offer a new method for proving the following result that was established in [19]. For non-complete graphs,

$$\lambda_n \geq \frac{n+1}{n-1},$$

with equality if and only if the graph is either the complete graph with one edge removed, or a graph given by two copies of the complete graph on $\frac{n-1}{2}$ vertices, joined by a vertex of degree $n-1$. In contrast to [19, Theorem 3.1], our proof methods are completely different and allow for an extension of this estimate in terms of the minimal vertex degree. In particular, we prove that, for a graph with minimum vertex degree $d_{\min} \leq \frac{n-1}{2}$, λ_n is bounded below by a function depending only on n and on $d_{\min}$.

(b) We define a new Cheeger-like constant and we use it for proving Cheeger-like inequalities that bound the largest eigenvalue.

(c) We define two normalized Laplace operators for hypergraphs that can be useful in the study of chemical reaction networks and we investigate some properties of their spectra.

(d) Finally, we prove some new results on spectral measures and spectral classes.

Keywords: Laplace operator, spectral graph theory, hypergraphs, Cheeger constant.

Contents

1 Introduction

1.1 Networks

Networks are everywhere around us. Their history and their variety are beautifully discussed in the popular science books of the brilliant physicist Albert-László Barabási, as for example [4] and [5]. Significant networks are, for instance:

- Social networks. These include professional ties and friendships. Interestingly, network science emerged in the 2000s as a consequence of the impact of two classical papers and one of them, written by Mark Granovetter in 1973, is about social networks [26]. The other one is the 1959 paper by Paul Erdős and Alfréd Rényi in which random graphs have been defined for the first time [22].

- Chemical reaction networks. The first access to them has been made possible in the 1990s, when scientists collected the list of chemical reactions in a cell (which were already known) in central databases.

- Metabolic networks. These are comparable to the chemical reaction networks, as we shall also see in Chapter 5. They have important applications in medicine: for example, they help identifying the right drugs for both humans and bacteria.

- Neural networks. Neurons are cells in the nervous system whose function is to receive and transmit information; capturing the connections between them is therefore fundamental in order to understand and describe them. The first mathematical model of neural networks has been introduced by the logician Walter Pitts and the neuroscientist Warren McCulloch in 1943, in a paper titled *A logical calculus of the ideas immanent in nervous activity* [49].

- Communication networks, which describe the interactions between communication devices. An interesting study of this network is the one proposed by Isella *et al.* in the paper *What's in a crowd? Analysis of face-to-face behavioral networks* [33]. Here, the authors analyze the behaviors of a crowd by tracking face-to-face proximities between people. This is made possible by the conference budges that people are wearing, which contain Radio-Frequency Identification Devices.

- Epidemics networks. The first pandemic whose evolution have been predicted well in advance thanks to network science has been the 2009 H1N1 Pandemic [2].

- Terrorism networks, which are used for fighting terrorism.

- Scientific collaboration networks. In mathematics, these networks are used to compute the well-known *Erdős number* of a researcher, i.e. the distance from the famous mathematician Paul Erdős in terms of collaborations in mathematical publications [58]. Since Erdős has published more than 1.500 papers with more than 500 authors, it's not difficult to have a small Erdős number. The author of this dissertation, for example, has Erdős number ≤ 4.[1]

Graph and hypergraph theory offers the perfect mathematical models for networks, as we shall see in the next sections.

1.2 Graphs

Graph theory was born in 1736, when Leonard Euler solved the famous Königsberg bridges problem [24], but the word *graph* has been actually used for the first time in 1878 by James J. Sylvester [56, 23].
Graphs are a widely-used model for representing networks. For instance, Figure 1.1 shows part of the collaboration network of Paul Erdős, drawn by Ron Graham as a graph. Figure 1.2 shows the graph of Medieval trade routes.

We give some basic definitions and assumptions concerning graphs. We consider, in particular, unweighted and undirected graphs without loops, multiple edges and isolated vertices.

[1]This can be proved in two different ways. P. Erdős coauthored with Ronald L. Graham, who coauthored with Shing-Tung Yau, who coauthored with Jürgen Jost, who coauthored with R. Mulas. Analogously, P. Erdős coauthored with Persi Diaconis, who coauthored with Bernd Sturmfels, who coauthored with Ngoc M. Tran, who coauthored with R. Mulas.

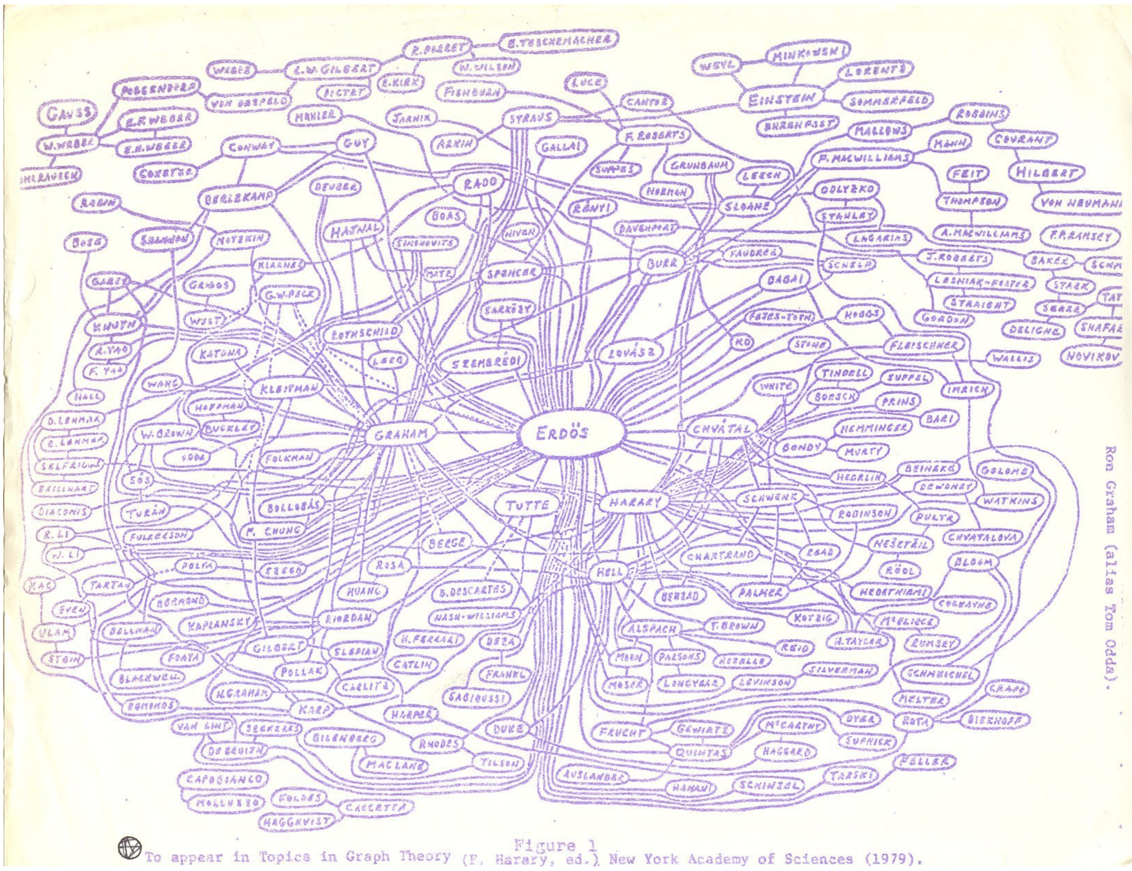


Figure 1.1: Part of the collaboration network of Paul Erdős, drawn by Ron Graham. Source: [21].

Definition 1.2.1. *A graph $\Gamma = (V, E)$ is a pair given by:*

- *A nonempty, finite set $V = V(\Gamma)$ called the vertex set, whose elements are called nodes or vertices;*
- *The edge set $E = E(\Gamma)$, whose elements are pairs $e = (v, w)$ of distinct vertices v and w called the endpoints of e.*

We assume, in particular, that the edges are unordered pairs. If $e = (v, w) \in E(\Gamma)$ we say that the vertices v and w are connected, or adjacent, or neighbors and we write $v \sim w$. Given a vertex v, we let $\mathcal{N}(v) \subset V(\Gamma)$ be the set of neighbors of v and we define the degree of v, denoted $\deg v$, as the cardinality of $\mathcal{N}(v)$. We assume that there are no vertices of degree zero.

Definition 1.2.2. *Two graphs $\Gamma = (V, E)$ and $\Gamma' = (V', E')$ are isomorphic if there exists a bijection $\varphi : V \to V'$ such that*

$$v \sim w \qquad \Longleftrightarrow \qquad \varphi(v) \sim \varphi(w).$$

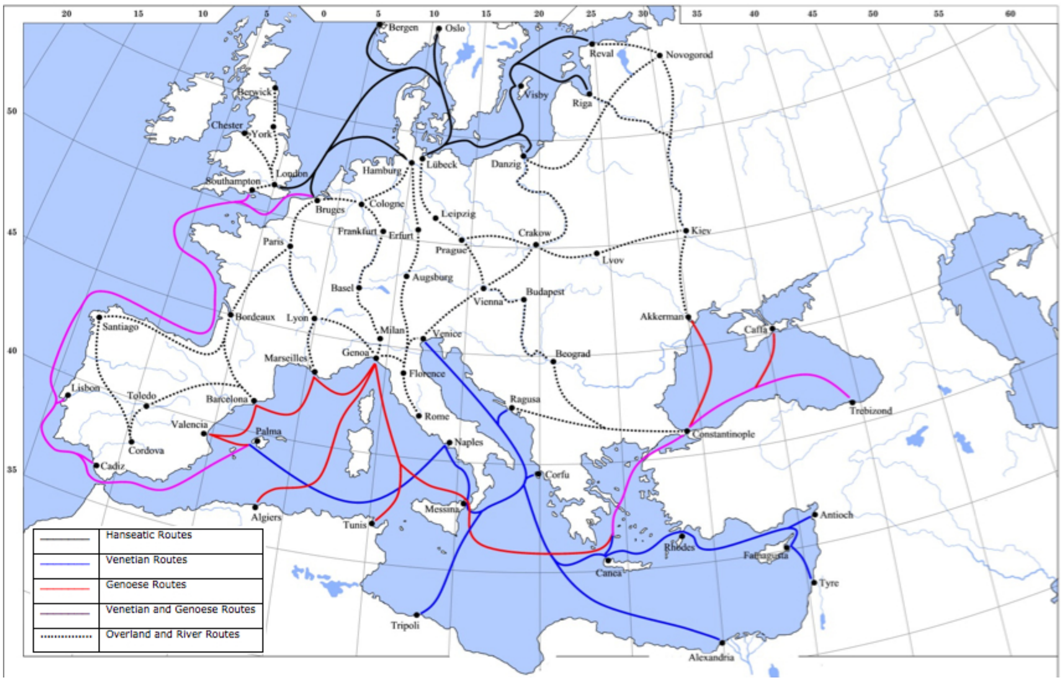


Figure 1.2: Graph of Medieval trade routes. Source: [21].

Here we consider graphs up to isomorphism.

Definition 1.2.3. *We say that a graph $\Gamma' = (V', E')$ is a subgraph of $\Gamma = (V, E)$, and we write $\Gamma' \subset \Gamma$, if $V' \subseteq V$ and $E \subseteq E'$.*

Example 1.2.4. *A graph on n nodes is said to be (Figure 1.3):*

- *Complete, denoted K_n, if $(v, w) \in E$ for every pair of distinct vertices $v, w \in V$.*

- *Biparite, if there exists a bipartition of the vertex set into two disjoint sets $V = V_2 \sqcup V_2$ such that each edge in E is between a vertex in V_1 and a vertex in V_2.*

- *Complete bipartite, if it is bipartite and there are all possible edges between V_1 and V_2, i.e.*

$$E = \{(v_1, v_2) : v_1 \in V_1 \text{ and } v_2 \in V_2\}.$$

  The complete bipartite graph with $|V_1| = n_1$ and $|V_2| = n_2$ is denoted K_{n_1,n_2}.

- *A star graph, denoted S_{n-1}, if it is the complete bipartite graph $K_{1,n-1}$.*

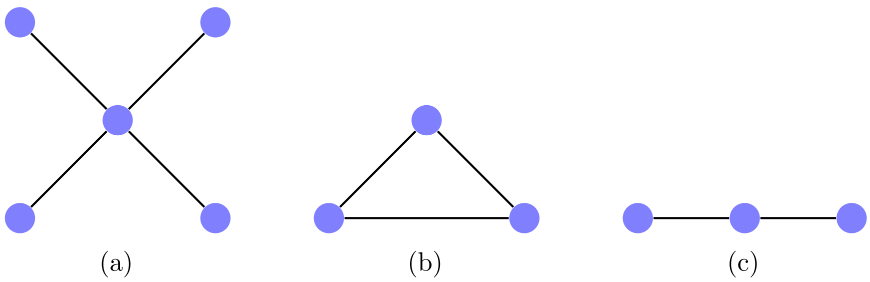


Figure 1.3: On the left: the star graph S_4, which is a particular case of complete bipartite graph. In the middle: the complete graph K_3, which coincides with the cycle C_3 and is a 2-regular graph. On the right: the path P_3.

- *d-regular, for some $d \in \mathbb{N}$ with $0 < d < n$, if all vertices have degree d.*
- *A petal graph, if one node has degree $n-1$ and all other nodes have degree 2. In particular, the petal graph on n nodes is well defined only if $n = 2m+1$ for some $m \geq 1$.*
- *A path (between v_1 and v_n), denoted P_n, if it is of the form*

$$V = \{v_1, \dots, v_n\}, \qquad E = \{(v_1, v_2), (v_2, v_3), \dots, (v_{n-1}, v_n)\}.$$

- *A cycle, denoted C_n, if it is of the form*

$$V = \{v_1, \dots, v_n\}, \qquad E = \{(v_1, v_2), (v_2, v_3), \dots, (v_{n-1}, v_n), (v_n, v_1)\}.$$

Definition 1.2.5. *A graph is connected if, for every pair of distinct vertices $v, w \in V$, there exists a path between v and w in Γ. A graph Γ that is non-connected is made by more than one connected subgraph, and these are called the connected components of Γ.*

Definition 1.2.6. *A directed graph is a graph where the edges $e = (v, w)$ are ordered pairs, called directed edges.*

While we shall not work with directed edges, we shall nevertheless work with *oriented* edges. That is, an edge e with endpoints v and w can carry two orientations, one going from v to w and the other in the opposite direction. We arbitrarily call the two orientations of a edge e $+$ and $-$. Analogously to differential forms in Riemannian geometry, see for instance [35], we shall consider functions γ from the set of oriented edges that satisfy

$$\gamma(e,-) = -\gamma(e,+), \tag{1.2.1}$$

that is, changing the orientation of e produces a minus sign. Importantly, neither of the two orientations that an edge carries plays a preferred role.

Definition 1.2.7. *To a graph $\Gamma = (V, E)$ with vertices $v_1, \ldots, v_n$, we associate the following matrices:*

- *We let $A = A(\Gamma)$ be the $n \times n$ adjacency matrix, defined by*

$$A_{ij} := \begin{cases} 1 & \textit{if } v_i \sim v_j \\ 0 & \textit{otherwise.} \end{cases}$$

- *We let $D = D(\Gamma)$ be the $n \times n$ diagonal degree matrix, defined by*

$$D_{ij} := \begin{cases} \deg v_i & \textit{if } i = j \\ 0 & \textit{otherwise.} \end{cases}$$

1.3 Simplicial complexes and hypergraphs

Simplicial complexes (Figure 1.4) are a generalization of graphs that can be used as a model for networks when a higher-order structure is needed [48]. While graphs have points and line segments as building blocks from a geometrical point of view, simplicial complexes are made by *simplices*: points, line segments, filled-in triangels, solid tetrahedra, and their higher-dimensional analogoues [44]. In applications this means that, while graphs identify pairwise-connections between elements, simplicial complexes identify communities.

An even more general notion is the one of a hypergraph (Figure 1.5), where one doesn't have a precise geometrical and topological representation (as the one that one can have with simplicial complexes), to the advantage of a more detailed description of the communities connections.
Both hypergraphs and simplicial complexes are widely used for instance when modelling neural networks (see for instance [51, 18, 17, 10]). Here, communities indicate neurons that are active together at a given point in time.

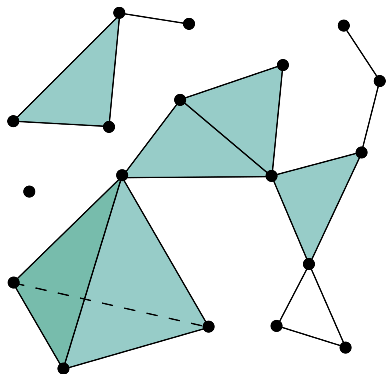


Figure 1.4: The geometrical representation of a simplicial complex. Source: [60].

Definition 1.3.1. *We define a hypergraph as a collection of nonempty finite sets. We call these sets hyperedges and we call their elements vertices.*

Definition 1.3.2. *We define a simplicial complex as a hypergraph $\mathcal{K}$ such that, if $k \in \mathcal{K}$, then every nonempty subset of k is in $\mathcal{K}$.*

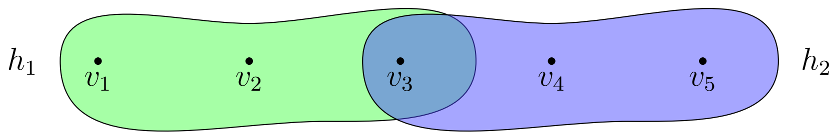


Figure 1.5: A hypergraph given by the hyperedges $h_1 = \{v_1, v_2, v_3\}$ and $h_2 = \{v_3, v_4, v_5\}$.

Observe that, while in a simplicial complex if $v_1, \ldots, v_n$ form a community, then also each subset of $\{v_1, \ldots, v_n\}$ forms a community, hypergraphs don't have this restriction. This is reasonable in many applications. Consider, for example, the following music collaboration network. We know that:

- John Cale, Nico and Lou Reed realized the album *Le Bataclan '72*;
- Nico never realized an album either alone with John Cale or alone with Lou Reed;
- John Cale and Lou Reed realized the album *Songs for Drella* together, without Nico.

A hypergraph could represent the sub-community made by John Cale and Lou Reed. A simplicial complex would contain all sub-communities of the largest community made by John Cale, Nico and Lou Reed.

In Chapter 5 we shall also refine the definition of a hypergraph in order to have a more precise model for chemical reaction networks.

1.4 Laplace operators

Pierre-Simon de Laplace was born in Normandy in 1749 and died in Paris at the age of seventy-seven, in the same month and year as Isaac Newton but exactly one hundred years later, and with similar last words[2] [11]. His family wanted him to become a priest and therefore Laplace, from the age of seven to the age of seventeen, attended a school of Benedictines. Then, he started to study theology at the University of Caen, but three years later he left the career that he had been preparing since childhood, because he fell in love with mathematics. His destiny was not the one that his family was planning for him. His destiny was to become a great mathematician and astronomer. His Laplace operator is the one defined for functions on Euclidean spaces,

$$\Delta f := \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \ldots + \frac{\partial^2 f}{\partial x_n^2}, \tag{1.4.1}$$

which can be generalized for Riemannian manifolds. Laplace worked on this operator in particular for studying the Laplace equation,

$$\Delta f = 0, \tag{1.4.2}$$

whose solutions are called harmonic functions.

Can one determine the shape of an object by listening to its vibrations? – this question has been first asked in 1882 by the physicist Arthur Schuster [9] and it can we reformulated as follows: *Can one reconstruct the shape of a mathematical drum form the eigenvalues of its Laplace operator?*. It became famous in 1966, when Mark Kac published the paper *Can one hear the shape of a drum?* [40] and the answer has been given in 1992 by Carolyn Gordon, David L. Webb and Scott Wolper in a paper titled *One cannot hear the shape of a drum* [25]. In other

[2]Issac Newton's last words in March 1727 were: *The great ocean of truth lay all undiscovered before me.* The last words of Laplace in March 1827 were: *What we know is not much, what we do not know is immense.*

words, one cannot reconstruct the exact shape of an object from the eigenvalues of the Laplacian. Nevertheless, one can infer important information. In 1911, Weyl proved an asymptotic formula for the eigenvalues of a compact Riemmanian manifold that depends on the volume of the manifold [59, 50]. In 1953, Minakshisundaram showed that for a compact Riemmanian manifold without boundary one can hear the dimension, the volume and the total scalar curvature [9, 50].

A discrete version of the Laplace operator in (1.4.1) has been used for the first time by Kirchhoff in 1847, for the study of electrical networks [41, 34]. This is the reason why the non-normalized Laplacian matrix of a graph Γ,

$$K := D - A,$$

is also called the Kirchhoff matrix.
In 1992, Fan Chung [16] introduced the first normalized version of K, the symmetric normalized Laplacian

$$\mathcal{L} := \mathrm{Id} - D^{-1/2} A D^{-1/2}$$

associated to a graph on n nodes, where Id is the $n \times n$ identity matrix.
In this thesis we choose to study the non-symmetric, normalized Laplacian

$$L := \mathrm{Id} - D^{-1} A,$$

also called the normalized random walk Laplacian because of its connection with random walks [53]. From now on, we shall simply call it the Laplacian. It's easy to check that

$$L = D^{-1/2} \mathcal{L} D^{1/2},$$

therefore the matrices L and $\mathcal{L}$ are similar, which implies that they have the same spectrum, i.e. the same eigenvalues counted with multiplicity. For a graph on n vertices these are exactly n and, as we shall see in Chapter 2, they are real and non-negative. We arrange them as

$$\lambda_1 \leq \ldots \leq \lambda_n$$

and we call them the graph eigenvalues. As already pointed out by Chung when she first introduced $\mathcal{L}$, two graphs cannot always be distinguished by their spectra (in other words, *we cannot hear the isomorphism class of a graph*), but the spectrum reveals some important properties. *Is the graph bipartite? Is it complete? How many connected components does it have?* – as we shall see, these are all questions that can be answered using the spectrum of the Laplace operator, so even if it does not distinguish the details of graphs, it does partition them into important families. We say, in particular, that two graphs are isospectral if they have the same spectrum. Since, furthermore, the computation of the eigenvalues can be performed with tools from linear algebra, studying the spectrum of these Laplacians is a very common tool in graph theory and data analytics.

1.4.1 Graph spectrum

> *Just as astronomers study stellar spectra to determine the make-up of distant stars, one of the main goals in graph theory is to deduce the principal properties and structure of a graph from its graph spectrum.*
>
> *(Fan Chung)*

We now summarize some of the basic results on the spectrum of the Laplace operator presented by Chung. We shall see more details in Chapter 2. We start with the observation that, since $\mathcal{L}$ is symmetric, its eigenvalues are real and non-negative.

Lemma 1.4.1. *[16, Lemma 1.7] For a graph Γ on n nodes, we have that:*

- $\lambda_1 = 0$ *and the multiplicity of the eigenvalue* 0 *equals the number of connected components of* Γ.
- *The largest eigenvalue is such that*

$$\lambda_n \geq \frac{n}{n-1}$$

  with equality if and only if Γ *is complete, and*

$$\lambda_n \leq 2$$

  with equality if and only if a connected component of Γ *is bipartite.*

- *The spectrum of a graph is the union of the spectra of its connected components.*

For connected graphs, the first non-zero eigenvalue λ_2 is controlled both above and below by the Cheeger constant h, a quantity that measures how difficult it is to partition the vertex set into two disjoint sets V_1 and V_2 such that the number of edges between V_1 and V_2 is as small as possible and such that the *volume* of both V_1 and V_2, i.e. the sum of the degrees of their vertices, is as big as possible. In particular, the Cheeger constant is defined as

$$h := \min_S \frac{|E(S,\bar{S})|}{\min\{\text{vol}(S), \text{vol}(\bar{S})\}}$$

where, given $\emptyset \neq S \subsetneq V$, $\bar{S} := V \setminus S$, $|E(S,\bar{S})|$ denotes the number of edges with one endpoint in S and the other in $\bar{S}$, and $\text{vol}(S) := \sum_{v\in S} \deg(v)$.
The following well-known theorem [16, Lemma 2.1 and Theorem 2.2] gives two important bounds for λ_2 in terms of h.

Theorem 1.4.2. *For every connected graph,*

$$\frac{1}{2}h^2 \leq \lambda_2 \leq 2h.$$

In particular, because of the definition of h, these bound imply that λ_2 estimates the coherence of the graph, that is, how different it is from a disconnected one.
Here is another historical note [36]. The graph Cheeger constant has been first introduced by the Hungarian mathematicians George Polya and Gábor Szegő in 1951 [54] and it's the discrete analogue of the constant that has been defined in Riemannian geometry by Jeff Cheeger in 1970 [15]. The graph Cheeger inequalities in Theorem 1.4.2 are the discrete analogue of Cheeger estimates on Riemannian manifolds and they have been proved by Jozef Dodziuk [20] and Alon–Milman [1] between 1984 and 1985.

1.5 Overview of the dissertation

This thesis is organized as follows.
In Chapter 2 we discuss more details about the spectrum of the Laplace operator, as preliminaries to our work. We present an alternative construction for the Laplacian which can be seen as a self-adjoint operator, we define the edge-Laplace operator that shares the same non-zero spectrum with L, and we discuss the min-max principle: a very powerful tool for getting insights about the graph eigenvalues.

In Chapter 3 we offer a new method for proving the following result that was established in [19]. For non-connected graphs,

$$\lambda_n \geq \frac{n+1}{n-1},$$

with equality if and only if the graph is either the complete graph with one edge removed, or a graph given by two copies of the complete graph on $\frac{n-1}{2}$ vertices, joined by a vertex of degree $n-1$. In contrast to [19, Theorem 3.1], our proof methods are completely different and allow for an extension of this estimate in terms of the minimal vertex degree. In particular, we show that for a graph with minimum vertex degree $d_{\min} \leq \frac{n-1}{2}$,

$$\lambda_n \geq 1 + \frac{1}{\sqrt{d_{\min}(n-1-d_{\min})}}.$$

In Chapter 4 we define a new Cheeger-like constant and we use it for proving Cheeger-like inequalities that bound the largest eigenvalue.
In Chapter 5 we define a new class of hypergraphs, that we call *chemical hypergraphs*, as a model for chemical reaction networks. We define two normalized Laplace operators on chemical hypergraphs that generalize the Laplace operator L and the edge-Laplacian, and we investigate some properties of their spectra.
In Chapter 6 we prove some new results on spectral measures and spectral classes.

1.6 Publications and preprints

The work of this thesis is based on the following publications and preprints.

- Chapter 3:
  J. Jost, R. Mulas, and F. Münch. *Communications in Mathematics and Statistics,* accepted. [39]

- Chapter 4:
  J. Jost and R. Mulas. Cheeger-like inequalities for the largest eigenvalue of the graph Laplace Operator. *arXiv:1910.12233.* [37]

- Chapter 5:
  J. Jost and R. Mulas. Hypergraph Laplace operators for chemical reaction networks. *Advances in Mathematics*, 351:870–896, 2019. [38]

- Chapter 6:
  A. Lerario and R. Mulas. Random geometric complexes and graphs on Riemannian manifolds in the thermodynamic limit. *Discrete & Computational Geometry,* accepted. [45]

2 Preliminaries

In this chapter we shall investigate some more known results about spectral graph theory, as preliminaries to our new results. For general references, the reader is referred to [31] and [16]. We fix for the rest of the chapter a graph $\Gamma = (V, E)$ on n nodes and we introduce the following notation. Given a set X, let $C(X)$ be the vector space of functions $f : X \to \mathbb{R}$. Then,

$$L = \mathrm{Id} - D^{-1}A$$

can be seen as $L : C(V) \to C(V)$ such that, for $v \in V$,

$$Lf(v) = f(v) - \frac{1}{\deg v} \sum_{w \sim v} f(w) = \sum_{w \sim v} \frac{f(v) - f(w)}{\deg v}.$$

Therefore, L computes the difference between the value of a function f at a vertex v and the average value of f at the neighbors of v.

2.1 Alternative construction of the Laplace operator

We discuss, now, an alternative way of constructing L.

We shall:

(i) Give weight one to edges and give weight $\deg v$ to each vertex v;

(ii) Define a scalar product for functions defined on edges and a scalar product for functions defined on vertices, based on the weights we gave;

(iii) Define the boundary operator for functions defined on the vertex set;

(iv) Consider the coboundary operator based on the scalar products we defined;

(v) Define the Laplace operator as a composition of the boundary and the coboundary operator.

Definition 2.1.1 (Scalar product for functions defined on oriented edges)**.** *Given $\omega, \gamma : E \to \mathbb{R}$, let*

$$\langle \omega, \gamma \rangle_E := \sum_{e \in E} \omega(e) \cdot \gamma(e).$$

Definition 2.1.2 (Scalar product for functions defined on vertices)**.** *Given $f, g : V \to \mathbb{R}$, let*

$$\langle f, g \rangle := \sum_{v \in V} \deg v \cdot f(v) \cdot g(v).$$

Definition 2.1.3 (Boundary operator for functions defined on vertices)**.** *Given $f : V \to \mathbb{R}$ and $e = (v, w)$ with an orientation such that v is the input and w is the output, let*

$$\delta f(e) := f(v) - f(w).$$

Remark 2.1.4. *Note that*

$$\delta : C(V) \longrightarrow C(E)$$

where the functions in $C(E)$ are always supposed to satisfy (1.2.1)*. In particular, δf also satisfies* (1.2.1)*.*

Definition 2.1.5 (Coboundary operator)**.** *Let*

$$\delta^* : C(E) \longrightarrow C(V)$$

be defined as

$$\delta^*(\gamma)(v) := \frac{\sum_{e_{in}: v \ input} \gamma(e_{in}) - \sum_{e_{out}: v \ output} \gamma(e_{out})}{\deg v}.$$

Remark 2.1.6. *In homology theory, the boundary operators act from higher to lower dimension. In cohomology theory, the boundary operators are the dual of the homology boundary operators and they act from lower to higher dimension. Therefore, the choice of the definition for boundary and coboundary operator also in our case depends on the point of view that one chooses to adopt.*

Lemma 2.1.7. *δ^* satisfies $\langle \delta f, \gamma \rangle_E = \langle f, \delta^* \gamma \rangle$, therefore it is the (unique) adjoint operator of δ.*

We shall see the proof of Lemma 2.1.7 directly in Chapter 5, when proving a more general version of it (Lemma 5.2.6). We now prove the following characterization of L.

Lemma 2.1.8. *The Laplace operator can be written as*

$$L = \delta^* \circ \delta.$$

Proof. For each $f \in C(V)$ and $v \in V$, we have that

$$\delta^*(\delta f)(v) = \frac{\sum_{e_{\text{in}}:v \text{ input}} \delta f(e_{\text{in}}) - \sum_{e_{\text{out}}:v \text{ output}} \delta f(e_{\text{out}})}{\deg v}$$

$$= \frac{\sum_{e_{\text{in}}=(v,w'):v \text{ input}} \Big(f(v) - f(w') \Big) - \sum_{e_{\text{out}}=(v,\hat{w}):v \text{ output}} \Big(f(\hat{w}) - f(v) \Big)}{\deg v}$$

$$= \sum_{w \sim v} \frac{f(v) - f(w)}{\deg v}$$

$$= Lf(v).$$

□

2.2 Edge-Laplacian

We have proved that

$$L = \delta^* \circ \delta.$$

We now define a new operator, the edge-Laplacian, as

$$L^E := \delta \circ \delta^*.$$

We first observe that:

- L and L^E are the two compositions of δ and δ^*, which are adjoint to each other. Therefore they are both self-adjoint, which implies that their eigenvalues are real.
- L and L^E are non-negative operators. In fact, given $f : V \to \mathbb{R}$,

  $$\langle Lf, f \rangle = \langle \delta^* \delta f, f \rangle = \langle \delta f, \delta f \rangle_E \geq 0.$$

  Analogously, for $\gamma : E \to \mathbb{R}$,

  $$\langle L^E \gamma, \gamma \rangle_E = \langle \delta \delta^* \gamma, \gamma \rangle_E = \langle \delta^* \gamma, \delta^* \gamma \rangle \geq 0.$$

  This implies that the eigenvalues of L and L^E are non-negative.

We prove, now, that L and L^E have the same non-zero spectrum.

Lemma 2.2.1. *If A and B are linear operators, then the non-zero eigenvalues of AB and BA are the same. In particular, if μ is a non-zero eigenvalue of AB for an eigenvector v, then μ is an eigenvalue of BA for the eigenvector Bv.*

Proof. Let μ be a non-zero eigenvalue of AB for a non-zero eigenvector v. Then

$$\mu Bv = B\mu v = B(ABv) = (BA)Bv.$$

Therefore, μ is an eigenvalue of BA for the eigenvector Bv. □

Corollary 2.2.2. *The non-zero eigenvalues of L and L^E are the same. In particular, if f is an eigenfunction of L with eigenvalue $\lambda \neq 0$, then δf is an eigenfunction of L^E with eigenvalue λ; if γ is an eigenfunction of L^E with eigenvalue $\lambda' \neq 0$, then $\delta^* \gamma$ is an eigenfunction of L with eigenvalue λ'.*

We therefore have three alternative ways to control or estimate the non-vanishing graph eigenvalues: through L, through $\mathcal{L}$ or through L^E.
As another important consequence of Corollary 2.2.2, the two operators only differ in the multiplicity of the eigenvalue 0. Let m_V and m_E be the multiplicity of the eigenvalue 0 of L and L^E, respectively. Then Corollary 2.2.2 implies

Corollary 2.2.3.

$$m_V - m_E = |V| - |E|. \tag{2.2.1}$$

Interestingly, $|V| - |E|$ is the Euler characteristic of the graph, that can also be written in terms of the Betti numbers as $\beta_0 - \beta_1$. Therefore (2.2.1) shows that L and L^E together capture this important topological invariant.

The above construction of L and L^E with boundary and coboundary operator has been generalized for simplicial complexes by Danijela Horak and Jürgen Jost in 2013 [31]. In particular, while in the case of graphs we can see L as a zero-dimensional operator and L^E as a one-dimensional operator, in [31] they define a k-dimensional Laplace operator for each k up to the dimension of the simplicial complex. In Chapter 5 we shall generalize the construction for graphs to the case of hypergraphs. Since, for general hypergraphs, higher dimensions are not well-defined, our new construction shall not generalize the one for simplicial complexes.

2.3 Min-max principle

In this section, we will apply the following theorem [36] in order to get more insight about the spectra of L and L^E.

Theorem 2.3.1 (Courant–Fischer–Weyl min-max principle)**.** *Let W be an n-dimensional vector space with a positive definite scalar product $(.,.)$. Let $\mathcal{W}_k$ be the family of all k-dimensional subspaces of W. Let $A : W \to W$ be a self-adjoint linear operator. Then the eigenvalues $\lambda_1 \leq \ldots \leq \lambda_n$ of A can be obtained by*

$$\lambda_k = \min_{W_k \in \mathcal{W}_k} \max_{g(\neq 0) \in W_k} \frac{(Ag, g)}{(g, g)} = \max_{W_{n-k+1} \in \mathcal{W}_{n-k+1}} \min_{g(\neq 0) \in W_{n-k+1}} \frac{(Ag, g)}{(g, g)}.$$

The vectors g_k realizing such a min-max or max-min then are corresponding eigenvectors, and the min-max spaces $\mathcal{W}_k$ are spanned by the eigenvectors for the eigenvalues $\lambda_1, \ldots, \lambda_k$, and analogously, the max-min spaces $\mathcal{W}_{n-k+1}$ are spanned by the eigenvectors for the eigenvalues $\lambda_k, \ldots, \lambda_n$. Thus, we also have

$$\lambda_k = \min_{g \in W, (g, g_j) = 0 \text{ for } j = 1, \ldots, k-1} \frac{(Ag, g)}{(g, g)} = \max_{g \in V, (g, g_l) = 0 \text{ for } l = k+1, \ldots, n} \frac{(Ag, g)}{(g, g)}. \tag{2.3.1}$$

In particular,

$$\lambda_1 = \min_{g \in W} \frac{(Ag, g)}{(g, g)}, \qquad \lambda_n = \max_{g \in W} \frac{(Ag, g)}{(g, g)}.$$

Definition 2.3.2. $\frac{(Ag,g)}{(g,g)}$ *is called the* Rayleigh quotient *of g.*

Remark 2.3.3. *Without loss of generality, we may assume $(g, g) = 1$ in (2.3.1).*

As a consequence of the min-max principle, we can write the eigenvalues of L and L^E as the *min-max* of the Rayleigh quotients

$$\frac{\langle Lf, f \rangle}{\langle f, f \rangle} = \frac{\sum_{v \sim w} \Big(f(v) - f(w) \Big)^2}{\sum_{v \in V} \deg v \cdot f(v)^2}$$

and

$$\frac{\langle L^E \gamma, \gamma \rangle_E}{\langle \gamma, \gamma \rangle_E} = \frac{\sum_{v \in V} \frac{1}{\deg v} \cdot \Big(\sum_{e_{\text{in}} : v \text{ input}} \gamma(e_{\text{in}}) - \sum_{e_{\text{out}} : v \text{ output}} \gamma(e_{\text{out}}) \Big)^2}{\sum_{e \in E} \gamma(e)^2}.$$

In particular, since it's clear that

$$\frac{\langle Lf, f \rangle}{\langle f, f \rangle} = 0 \tag{2.3.2}$$

if and only if f is constant on each connected component of the graph, by the min-max principle we have that the multiplicity of the eigenvalue 0 for L equals the number of connected components of the graph, i.e.

$$m_V = \beta_0. \tag{2.3.3}$$

By (2.2.1), this also implies that

$$m_E = \beta_1. \tag{2.3.4}$$

Because of the analogy between (2.3.2) and the Laplace equation (1.4.2), the eigenfunctions of 0 for a graph are called harmonic functions. Furthermore, by (2.3.1), the eigenfunctions of the non-zero eigenvalues have to be orthogonal to the harmonic functions. Therefore, they have to satisfy

$$\sum_{v\in V} \deg v \cdot f(v) = 0. \tag{2.3.5}$$

Putting everything together, we have that for any graph we can write

$$\lambda_n = \max_{f:V\to\mathbb{R}} \frac{\sum_{v\sim w}\Big(f(v)-f(w)\Big)^2}{\sum_{v\in V}\deg v\cdot f(v)^2} \tag{2.3.6}$$

$$= \max_{\gamma:E\to\mathbb{R}} \frac{\sum_{v\in V}\frac{1}{\deg v}\cdot\Big(\sum_{e_{\text{in}}:v\text{ input}}\gamma(e_{\text{in}})-\sum_{e_{\text{out}}:v\text{ output}}\gamma(e_{\text{out}})\Big)^2}{\sum_{e\in E}\gamma(e)^2} \tag{2.3.7}$$

and, for connected graphs,

$$\lambda_2 = \min_{f:V\to\mathbb{R}\text{ s.t. }\sum_{v\in V}\deg v\cdot f(v)=0} \frac{\sum_{v\sim w}\Big(f(v)-f(w)\Big)^2}{\sum_{v\in V}\deg v\cdot f(v)^2} \tag{2.3.8}$$

$$= \min_{f:V\to\mathbb{R}\text{ non constant}}\ \max_{t\in\mathbb{R}} \frac{\sum_{v\sim w}\Big(f(v)-f(w)\Big)^2}{\sum_{v\in V}\deg v\cdot\big(f(v)-t\big)^2}. \tag{2.3.9}$$

3 Spectral gap of the largest eigenvalue

For a general graph on n vertices, we know that

$$\lambda_n \geq \frac{n}{n-1},$$

with equality if and only if the graph is complete. Naturally, the question arises: *What is the optimal a-priori estimate for λ_n for non-complete graphs?* Das and Sun [19] proved that for all non-complete graphs one has

$$\lambda_n \geq \frac{n+1}{n-1},$$

with equality if and only if the complement graph is a single edge or a complete bipartite graph with both parts of size $\frac{n-1}{2}$. In this chapter we offer a new method for proving these results, see the proofs of Theorem 3.1.1 and Theorem 3.2.1. Furthermore, we use this new method for showing that, for a graph with minimum vertex degree $d_{\min} \leq \frac{n-1}{2}$,

$$\lambda_n \geq 1 + \frac{1}{\sqrt{d_{\min}(n-1-d_{\min})}}.$$

These results are presented also in [39], a joint work with Jürgen Jost and Florentin Münch.

Before proving the main results of this chapter, we recall that given $v \in V$ we defined $\mathcal{N}(v) \subset V$ to be the set of neighbors of v, and we give two new definitions.

Definition 3.0.1. *Given $k \in \mathbb{N} \cup \{\infty\}$, let*

$$\mathcal{N}_k(v) := \{w \in V : d(v,w) = k\},$$

where

$$d(v,w) := \inf\{n \in \mathbb{N} \cup \{\infty\} : v = x_0 \sim \ldots \sim x_n = w\}$$

is the combinatorial graph distance.

Definition 3.0.2. *Given a graph $\Gamma = (V, E)$, we define its complement as $\overline{\Gamma} := (\overline{V}, \overline{E})$, where*

$$\overline{E} := \{(v, w) : v, w \in V, v \neq w \text{ and } (v, w) \notin E\}$$

and

$$\overline{V} := \{v \in V : \text{ there exists } w \in V \text{ s.t. } (v, w) \in \overline{E}\}.$$

3.1 Eigenvalue estimate

The following theorem was established in [19] and gives the optimal a-priori lower bound on λ_n for all non-complete graphs. In contrast to [19, Theorem 3.1], our proof methods are completely different and allow for an extension of this estimate in terms of the minimal vertex degree, see Section 3.3.

Theorem 3.1.1 ([19, Theorem 3.1])**.** *Let $\Gamma = (V, E)$ be a non-complete graph with n vertices.*

Then,

$$\lambda_n \geq \frac{n+1}{n-1}.$$

Proof. We first assume that Γ is connected. Since Γ is not complete, there exists a vertex v with $\deg v \leq n-2$. As the graph is connected, we have $\mathcal{N}_2(v) \neq \emptyset$. Let $w \in \mathcal{N}_2(v)$. Then, $\deg w \leq n-2$ as v and w are not adjacent. Moreover, $\mathcal{N}(v) \cap \mathcal{N}(w) \neq \emptyset$.

We write $A := |\mathcal{N}(v) \cap \mathcal{N}(w)|$. We aim to find a function f with $\langle Lf, f\rangle \geq \frac{n+1}{n-1}\langle f, f\rangle$. To do so, it is convenient to choose f in such a way that $Lf = \frac{n+1}{n-1}f$ in v and w. Particularly, let $f : V \to \mathbb{R}$ be given by

$$f(x) := \begin{cases} -1 & : x \in \mathcal{N}(v) \cap \mathcal{N}(w), \\ \frac{n-1}{2}\frac{A}{\deg v} & : x = v, \\ \frac{n-1}{2}\frac{A}{\deg w} & : x = w, \\ 0 & : \text{otherwise}. \end{cases}$$

We observe

$$\deg v \cdot Lf(v) = \deg v \cdot f(v) + A = \frac{n+1}{2}A$$

and thus, $Lf(v) = \frac{n+1}{n-1}f(v)$. Similarly, $Lf(w) = \frac{n+1}{n-1}f(w)$. We now claim that $-Lf(x) \geq \frac{n+1}{n-1}$ for all $x \in \mathcal{N}(v) \cap \mathcal{N}(w)$. We observe $A \geq 1 \vee (\deg v + \deg w + 2 - n)$

where $\vee$ denotes the maximum, and we calculate

$$-Lf(x) = \frac{\deg x - |\mathcal{N}(x) \cap \mathcal{N}(v) \cap \mathcal{N}(w)| + f(v) + f(w)}{\deg x} \tag{3.1.1}$$
$$\geq 1 + \frac{1 - A + f(v) + f(w)}{\deg x}.$$

As $f(v) + f(w) \geq A$, we can use $\deg x \leq n-1$ and continue

$$\frac{1 - A + f(v) + f(w)}{\deg x} \geq \frac{1-A}{n-1} + \frac{A}{2\deg v} + \frac{A}{2 \deg w} \tag{3.1.2}$$
$$= \frac{1}{n-1} + A\left(\frac{1}{2\deg v} + \frac{1}{2\deg w} - \frac{1}{n-1}\right).$$

Since $\deg v \leq n-2$ and $\deg w \leq n-2$, we see that the term in brackets is positive and thus,

$$A\left(\frac{1}{2\deg v} + \frac{1}{2\deg w} - \frac{1}{n-1}\right) \geq \tag{3.1.3}$$
$$\geq [1 \vee (\deg v + \deg w + 2 - n)]\left(\frac{1}{2\deg v} + \frac{1}{2\deg w} - \frac{1}{n-1}\right).$$

We write $D := (\deg v + \deg w)/2$, and by the harmonic-arithmetic mean estimate, we have $\frac{1}{2\deg v} + \frac{1}{2\deg w} \geq \frac{1}{D}$ and thus,

$$A\left(\frac{1}{2\deg v} + \frac{1}{2\deg w} - \frac{1}{n-1}\right) \geq [1 \vee (2D + 2 - n)]\left(\frac{1}{D} - \frac{1}{n-1}\right). \tag{3.1.4}$$

We aim to show that the latter term is at least $\frac{1}{n-1}$ which, by multiplying with $D(n-1)$ and subtracting D, is equivalent to

$$[1 \vee (2D + 2 - n)](n - 1 - D) - D \geq 0. \tag{3.1.5}$$

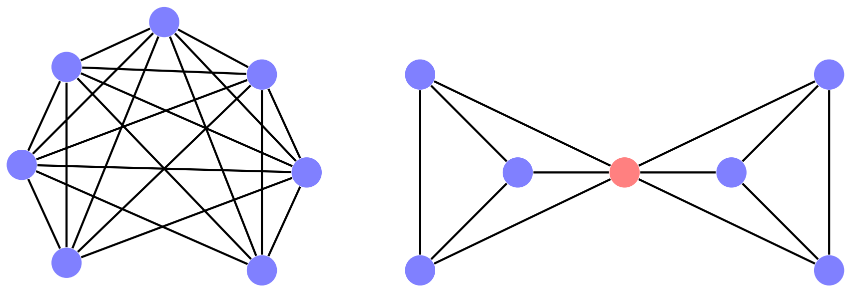
If $D \leq \frac{n-1}{2}$, then the maximum equals 1 and the inequality follows immediately. If $D \geq \frac{n-1}{2}$, then we can discard the "$1\vee$", and so the left hand side becomes a concave quadratic polynomial in D with its zero points in $D = n-2$ and $D = \frac{n-1}{2}$. Thus, the inequality (3.1.5) holds true for all D between the zero points. Moreover by assumption, D has to be between the zero points which proves the claim that $-Lf(x) \geq \frac{n+1}{n-1}$ for all $x \in \mathcal{N}(v) \cap \mathcal{N}(w)$. Particularly, this shows that $fLf \geq \frac{n+1}{n-1}f^2$. Integrating proves the claim of the theorem for all connected graphs. For non-connected graphs, the smallest connected component has at most $\frac{n}{2}$ vertices, and by the usual estimate, we get

$$\lambda_n \geq \frac{n/2}{n/2 - 1} = \frac{n}{n-2} > \frac{n+1}{n-1}$$

which proves the theorem for non-connected graphs. $\square$

3.2 Rigidity

We now show that Theorem 3.1.1 gives the optimal bound for non-connected graphs and we characterize equality in the eigenvalue estimate which can be attained only for two families of graphs (Figure 3.1). One family is made by complete graphs with one edge removed. The graphs in the other family are surprisingly significantly different. They can be seen as two copies of a complete graph which are joined by a single vertex. Again, our proof methods differ widely from [19].



(a) (b)

Figure 3.1: For $n = 7$, these are the two graphs in Theorem 3.2.1. The graph on the left is the complete graph K_7 with one edge removed. The graph on the right is made by two copies of the complete graph K_3, joined by the red vertex in the middle.

Theorem 3.2.1 ([19, Theorem 3.1]). *Let $\Gamma = (V, E)$ be a graph with n vertices. The following are equivalent:*

(i) $\lambda_n = \frac{n+1}{n-1}$,

(ii) The complement graph of Γ is a single edge or a complete bipartite graph with both parts of size $\frac{n-1}{2}$.

Proof. We first prove $(i) \Rightarrow (ii)$. We first note that Γ is non-complete but connected by the proof of Theorem 3.1.1. Thus, all inequalities in the proof of Theorem 3.1.1 must be equalities. Let $v \not\sim w$ with $\mathcal{N}(v) \cap \mathcal{N}(w) \neq \emptyset$. By equality in (3.1.1), all vertices within $\mathcal{N}(v) \cap \mathcal{N}(w)$ must be adjacent. By equality in (3.1.2), all vertices of $\mathcal{N}(v) \cap \mathcal{N}(w)$ must have degree $n-1$. By equality in (3.1.3), obtain

$$|\mathcal{N}(v) \cap \mathcal{N}(w)| = (\deg v + \deg w + 2 - n) \vee 1.$$

By equality in (3.1.4), we obtain $\deg v = \deg w = D$. Finally by equality in (3.1.5), we see that

$$D \in \left\{ \frac{n-1}{2}, n-2 \right\}.$$

We first show that if $D = n-2$, then the complement graph is a single edge. If $\deg v = \deg w = D = n-2$, then we get $|\mathcal{N}(v) \cap \mathcal{N}(w)| = n-2$. Since all vertices within $\mathcal{N}(v) \cap \mathcal{N}(w)$ are adjacent, we see that the only missing edge is the one from v to w which shows that the complement graph is a single edge.

Now we assume $D = \frac{n-1}{2}$. Then, $A = |\mathcal{N}(v) \cap \mathcal{N}(w)| = 1$ and we can write $\mathcal{N}(v) \cap \mathcal{N}(w) = \{x\}$. We recall that $\deg x = n-1$. We now specify the parts of the bipartite graph which we want to be the complement graph. One part is

$$P_v := \{v\} \cup \mathcal{N}(v) \setminus \{x\}$$

and similarly, $P_w := \{w\} \cup \mathcal{N}(w) \setminus \{x\}$. Let $\widetilde{v} \in P_v$ and $\widetilde{w} \in P_w$. Then $d(v, \widetilde{w}) = d(w, \widetilde{v}) = 2$ as x is adjacent to every other vertex. By applying the above arguments to the pair $(v, \widetilde{w})$, we see that $\mathcal{N}(v) \cap \mathcal{N}(\widetilde{w}) = \{x\}$. Particularly, $\widetilde{v} \not\sim \widetilde{w}$. Moreover, we have $\deg \widetilde{w} = \deg v = \frac{n-1}{2}$ and similarly, $\deg \widetilde{v} = \frac{n-1}{2}$. By a counting argument, this shows that every $\widetilde{v} \in P_v$ is adjacent to every vertex not belonging to P_w. An analogous statements holds for all $\widetilde{w} \in P_w$. Putting everything together, we see that the complement graph of Γ is precisely the complete bipartite graph with the parts P_v and P_w. This finishes the case distinction and thus, the proof of the implication $(i) \Rightarrow (ii)$ is complete.

We finally prove $(ii) \Rightarrow (i)$. We start with the case that the complement graph is the complete bipartite graph. Let the parts be P and Q. Then, $\phi := 1_P - 1_Q$ is eigenfunction to the eigenvalue $\frac{2}{n-1}$ and every function orthogonal to ϕ and 1 is eigenfunction to the eigenvalue $\frac{n+1}{n-1}$.

We end with the case that the complement graph is a single edge (v, w). Then, $\phi = 1_v - 1_w$ is eigenfunction to eigenvalue 1, and $\psi = -2 + (n+1)(1_v + 1_w)$ is eigenfunction to eigenvalue $\frac{n+1}{n-1}$. Every function orthogonal to ϕ, ψ and 1 is eigenfunction to the eigenvalue $\frac{n}{n+1}$.

This finishes the proof of $(ii) \Rightarrow (i)$ and thus, the proof of the theorem is complete. □

Remark 3.2.2. *In the second equality case in Theorem 3.2.1, for $n > 3$, the eigenvalue λ_n has multiplicity larger than 1. With the notation of the proof of that theorem, we can take any vertex $v' \in P_v$ and any vertex $w' \in P_w$ and a function that is 1 at v' and w', -1 at their single joint neighbor z, and 0 everywhere else.*

For $n = 5$, that is when we have two triangles sharing a single vertex z. We can also take a function that is 0 at z and assumes the values ± 1 on the two other vertices of each of the two triangles, to produce other eigenfunctions with eigenvalue $\frac{3}{2}$.

3.3 Lower bound using the minimum degree

We now use the same method as that in the proof of Theorem 3.1.1 in order to give a new lower bound to the largest eigenvalue in terms of the minimum vertex degree, provided this is at most $\frac{n-1}{2}$. To the best of our knowledge, this is the first known lower bound to λ_n in terms of the minimum degree. Li, Guo and Shiu [46] proved a bound in terms of the maximum degree. Namely, they have shown that for a graph with n vertices and m edges

$$\lambda_n \geq \frac{2m}{2m - \Delta},$$

where Δ is the maximum vertex degree.

Theorem 3.3.1. *Let $\Gamma = (V, E)$ be a graph with n vertices and let $d_{\min}$ be the minimum vertex degree of Γ. If $d_{\min} \leq \frac{n-1}{2}$, then*

$$\lambda_n \geq 1 + \frac{1}{\sqrt{d_{\min}(n - 1 - d_{\min})}}.$$

Proof. Let

$$\psi := \psi(n, d_{\min}) := 1 + \frac{1}{\sqrt{d_{\min}(n - 1 - d_{\min})}} \qquad \text{and} \qquad \eta := \frac{1}{\psi - 1}.$$

We proceed similarly to the proof of Theorem 3.1.1. We first assume that Γ is connected. Let v be a vertex of minimum degree, i.e. such that $\deg v = d_{\min}$. As the graph is connected, we have $N_2(v) \neq \emptyset$. Let $w \in N_2(v)$. Then, $\deg w \leq n - 2$ as v and w are not adjacent. Moreover, $N(v) \cap N(w) \neq \emptyset$.

We write $A := |N(v) \cap N(w)|$. We aim to find a function f with $\langle f, Lf \rangle \geq \psi \langle f, f \rangle$. To do so, it is convenient to choose f in such a way that $Lf = \psi f$ in v and w. Particularly, let $f : V \to \mathbb{R}$ be given by

$$f(x) := \begin{cases} -1 & : x \in N(v) \cap N(w), \\ \eta \cdot \frac{A}{\deg v} & : x = v, \\ \eta \cdot \frac{A}{\deg w} & : x = w, \\ 0 & : \text{otherwise.} \end{cases}$$

We observe
$$Lf(v) = (\eta+1)\cdot\frac{A}{\deg v} = \frac{\eta+1}{\eta}\cdot f(v) = \psi f(v)$$
and similarly, $Lf(w) = \psi f(w)$. We now claim that $-Lf(x) \geq \psi$ for all $x \in N(v)\cap N(w)$. We observe $A \geq 1 \vee (\deg v + \deg w + 2 - n)$ and we calculate
$$-Lf(x) = \frac{\deg x - |N(x)\cap N(v)\cap N(w)| + f(v) + f(w)}{\deg x} \geq 1 + \frac{1 - A + f(v) + f(w)}{\deg x}. \tag{3.3.1}$$

In order to proceed, we will use the following lemma which will be proven later independently.

Lemma 3.3.2. *We have*
$$\frac{\eta}{n-1}\cdot[1\vee(\deg v + \deg w + 2 - n)]\left(\frac{1}{\deg v} + \frac{1}{\deg w} - \frac{1}{\eta}\right) \geq \frac{1}{\eta} - \frac{1}{n-1}.$$

Applying the lemma and using $A \geq 1 \vee (\deg v + \deg w + 2 - n)$ gives
$$\begin{aligned}0 < \frac{1}{\eta} - \frac{1}{n-1} &\leq \frac{\eta}{n-1}\cdot[1\vee(\deg v + \deg w + 2 - n)]\left(\frac{1}{\deg v} + \frac{1}{\deg w} - \frac{1}{\eta}\right)\\ &\leq \frac{\eta A}{n-1}\cdot\left(\frac{1}{\deg v} + \frac{1}{\deg w} - \frac{1}{\eta}\right).\end{aligned}$$

Moving $\frac{1}{n-1}$ to the right hand side and using $\deg x \leq n-1$ gives
$$\begin{aligned}0 < \frac{1}{\eta} &\leq \frac{1}{n-1} + \frac{\eta A}{n-1}\cdot\left(\frac{1}{\deg v} + \frac{1}{\deg w} - \frac{1}{\eta}\right)\\ &\leq \frac{1}{\deg x} + \frac{\eta A}{\deg x}\cdot\left(\frac{1}{\deg v} + \frac{1}{\deg w} - \frac{1}{\eta}\right)\\ &= \frac{1 - A + f(v) + f(w)}{\deg x}\\ &\leq -Lf(x) - 1\end{aligned}$$

where we used (3.3.1) in the last estimate. Thus, $-Lf(x) \geq 1 + \frac{1}{\eta} = \psi$ for all $x \in N(v)\cap N(w)$. Integrating gives $\langle Lf, f\rangle \geq \psi\langle f, f\rangle$ which proves the theorem for all connected graphs. For a non-connected graph, we apply the theorem for the connected component containing v and use that ψ is decreasing in n. The proof of the theorem is now complete up to the proof of the lemma. □

Proof of the lemma. We consider two cases.

(a) Case 1: $\deg v + \deg w \leq n-1$. Then,

$$1 \vee (\deg v + \deg w + 2 - n) = 1$$

and

$$\frac{1}{\deg w} \geq \frac{1}{n-1-\deg v}.$$

Therefore

$$\frac{\eta}{n-1} \cdot \left(\frac{1}{\deg v} + \frac{1}{\deg w} - \frac{1}{\eta} \right) \geq \frac{\eta}{n-1} \cdot \left(\frac{1}{\deg v} + \frac{1}{n-1-\deg v} - \frac{1}{\eta} \right).$$

Now, we have that

$$\frac{\eta}{n-1} \cdot \left(\frac{1}{\deg v} + \frac{1}{n-1-\deg v} - \frac{1}{\eta} \right) \geq \frac{1}{\eta} - \frac{1}{n-1}$$

$$\Longleftrightarrow$$

$$\frac{\eta}{n-1} \cdot \left(\frac{1}{\deg v} + \frac{1}{n-1-\deg v} \right) \geq \frac{1}{\eta}$$

$$\Longleftrightarrow$$

$$\frac{1}{n-1} \cdot \left(\frac{1}{\deg v} + \frac{1}{n-1-\deg v} \right) \geq \frac{1}{\eta^2}.$$

This is true by definition of η and it is actually an equality.

(b) Case 2: $\deg v + \deg w > n-1$. Then,

$$1 \vee (\deg v + \deg w + 2 - n) = \deg v + \deg w + 2 - n \geq 2.$$

Therefore, it suffices to prove that

$$\frac{2\eta}{n-1} \cdot \left(\frac{1}{\deg v} + \frac{1}{\deg w} - \frac{1}{\eta} \right) \geq \frac{1}{\eta} - \frac{1}{n-1},$$

i.e. that

$$\frac{2\eta}{n-1} \cdot \left(\frac{1}{\deg v} + \frac{1}{\deg w} \right) \geq \frac{1}{\eta} + \frac{1}{n-1},$$

that can be re-written as

$$\frac{2}{n-1} \cdot \left(\frac{1}{\deg v} + \frac{1}{\deg w}\right) - \frac{1}{\eta^2} \geq \frac{1}{(n-1)\eta}.$$

In order to prove it, we use the fact that

$$\frac{1}{\eta^2} = \frac{1}{n-1} \cdot \left(\frac{1}{\deg v} + \frac{1}{n-1-\deg v}\right).$$

This implies that

$$\begin{aligned}
&\frac{2}{n-1} \cdot \left(\frac{1}{\deg v} + \frac{1}{\deg w}\right) - \frac{1}{\eta^2} \\
\geq &\frac{2}{n-1} \cdot \left(\frac{1}{\deg v} + \frac{1}{n-1}\right) - \frac{1}{n-1} \cdot \left(\frac{1}{\deg v} + \frac{1}{n-1-\deg v}\right) \\
= &\frac{1}{n-1} \cdot \left(\frac{1}{\deg v} + \frac{2}{n-1} - \frac{1}{n-1-\deg v}\right) \\
= &\frac{1}{n-1} \cdot \left(\frac{(n-1)(n-1-\deg v) + 2\deg v(n-1-\deg v) - \deg v(n-1)}{\deg v(n-1)(n-1-\deg v)}\right) \\
= &\frac{(n-1)^2 - \deg v(n-1) + 2\deg v(n-1) - 2\deg v^2 - \deg v(n-1)}{\deg v(n-1)^2(n-1-\deg v)} \\
= &\frac{(n-1)^2 - 2\deg v^2}{\deg v(n-1)^2(n-1-\deg v)}.
\end{aligned}$$

Therefore, the inequality that we want to prove becomes

$$\frac{(n-1)^2 - 2\deg v^2}{\deg v(n-1)^2(n-1-\deg v)} \geq \frac{1}{(n-1)\eta}$$

$$\iff$$

$$\frac{(n-1)^2 - 2\deg v^2}{\deg v(n-1)(n-1-\deg v)} \geq \frac{1}{\eta} = \frac{1}{\sqrt{\deg v(n-1-\deg v)}}$$

$$\iff$$

$$(n-1)^2 - 2\deg v^2 \geq (n-1)\sqrt{(n-1-\deg v) \cdot \deg v}.$$

Now, since we are assuming $\deg v \leq \frac{n-1}{2}$,

$$(n-1)^2 - 2\deg v^2 \geq \frac{(n-1)^2}{2}.$$

Also, by applying $\sqrt{ab} \leq \frac{a+b}{2}$,

$$(n-1)\sqrt{(n-1-\deg v)\cdot \deg v} \leq (n-1)\frac{n-1}{2} = \frac{(n-1)^2}{2}.$$

Therefore,

$$(n-1)^2 - 2\deg v^2 \geq (n-1)\sqrt{(n-1-\deg v)\cdot \deg v}.$$

Thus, the proof of the lemma is complete. □

Remark 3.3.3. *In the particular case of $d_{\min} = \frac{n-1}{2}$, Theorem 3.3.1 tells us that*

$$\lambda_n \geq \frac{n+1}{n-1}.$$

By the second part of Theorem 3.2.1 we know that this inequality is sharp.

4 Cheeger-like inequalities for the largest eigenvalue

Here we define a new Cheeger-like constant for graphs and we use it for proving Cheeger-like inequalities that bound the largest eigenvalue of the normalized Laplace operator. We also prove new general results of spectral graph theory that are useful in order to prove or discuss our main theorem. The results of this chapter are presented also in [37], a joint work with Jürgen Jost.

4.1 Motivation

For a fixed connected graph $\Gamma = (V, E)$ on n nodes, we have already seen in the introduction that the Cheeger constant is defined as

$$h := \min_S \frac{|E(S, \bar{S})|}{\min\{\mathrm{vol}(S), \mathrm{vol}(\bar{S})\}}$$

where, given $\emptyset \neq S \subsetneq V$, $\bar{S} := V \setminus S$, $|E(S, \bar{S})|$ denotes the number of edges with one endpoint in S and the other in $\bar{S}$, and $\mathrm{vol}(S) := \sum_{v \in S} \deg(v)$. We have also seen, in Theorem 1.4.2, that two important bounds for λ_2 in terms of h are given:

$$\frac{1}{2} h^2 \leq \lambda_2 \leq 2h.$$

Furthermore, recall that, as we have seen in Chapter 2, using the min-max Principle we can write

$$\begin{aligned}
\lambda_2 &= \min_{f: V \to \mathbb{R} \text{ s.t. } \sum_{v \in V} \deg v \cdot f(v) = 0} \frac{\sum_{v \sim w} \Big(f(v) - f(w) \Big)^2}{\sum_{v \in V} \deg v \cdot f(v)^2} \\
&= \min_{f: V \to \mathbb{R} \text{ non constant}} \max_{t \in \mathbb{R}} \frac{\sum_{v \sim w} \Big(f(v) - f(w) \Big)^2}{\sum_{v \in V} \deg v \cdot \big(f(v) - t \big)^2}.
\end{aligned}$$

The following theorem [16, Theorem 2.8 and Corollary 2.9] shows an interesting relation between h and λ_2 when, in the characterizations of λ_2 via the Rayleigh quotient, we replace the L_2-norm by the L_1-norm both in the numerator and denominator.

Theorem 4.1.1. *For every connected graph,*

$$h = \min_{f:V\to\mathbb{R} \text{ non constant}} \max_{t\in\mathbb{R}} \frac{\sum_{v\sim w}\left|f(v)-f(w)\right|}{\sum_{v\in V}\deg v\cdot\left|f(v)-t\right|}$$

and

$$\frac{1}{2}h \leq \min_{f:V\to\mathbb{R} \text{ s.t. } \sum_{v\in V}\deg v\cdot f(v)=0} \frac{\sum_{v\sim w}\left|f(v)-f(w)\right|}{\sum_{v\in V}\deg v\cdot\left|f(v)\right|} \leq h.$$

Remark 4.1.2. *Interestingly, the quantity*

$$\min_{f:V\to\mathbb{R} \text{ non constant}} \max_{t\in\mathbb{R}} \frac{\sum_{v\sim w}\left|f(v)-f(w)\right|}{\sum_{v\in V}\deg v\cdot\left|f(v)-t\right|}$$

that characterizes h in Theorem 4.1.1 is also equal to the second smallest eigenvalue of the 1-Laplacian *[13, 12, 30, 29].*

The goal of this chapter is to present an analogous study for the largest eigenvalue. In particular, for any given graph we introduce the new constant

$$Q := \max_{\text{edges }(v,w)}\left(\frac{1}{\deg v}+\frac{1}{\deg w}\right),$$

we prove that it can be characterized by writing λ_n using the Rayleigh quotient and then replacing the L_2-norm by the L_1-norm both in the numerator and denominator, and we prove that it controls the largest eigenvalue λ_n both above and below. Therefore, Q is an analogue of the Cheeger constant for the largest eigenvalue.

Note that in the literature there exists already a Cheeger-like constant that bounds the largest eigenvalue [8, 7]. It is defined as

$$\bar{h} := \max_{\text{partitions } V=V_1\sqcup V_2\sqcup V_3} \frac{|E(V_1,V_2)|}{\mathrm{vol}(V_1)+\mathrm{vol}(V_2)},$$

it is called the *dual Cheeger constant* and it's such that

$$2\bar{h} \leq \lambda_n \leq 1+\sqrt{1-(1-\bar{h})^2}.$$

Moreover, there are interesting results that show that h and $\bar{h}$ are actually related to each other. For the dual Cheeger constant, however, as far as we know there is no result analogous to Theorem 4.1.1. This gives us the motivation for defining the new constant Q that again bounds λ_n and, additionally, satisfies an analogue of Theorem 4.1.1.

4.2 Main results

For the rest of this chapter, we fix a graph $\Gamma = (V, E)$ on n vertices. Recall that, using the min-max Principle, we can write the largest eigenvalue associated to Γ as

$$\begin{aligned}\lambda_n &= \max_{f:V\to\mathbb{R}} \frac{\sum_{v\sim w}\Big(f(v)-f(w)\Big)^2}{\sum_{v\in V}\deg v\cdot f(v)^2}\\ &= \max_{\gamma:E\to\mathbb{R}} \frac{\sum_{v\in V}\frac{1}{\deg v}\cdot\Big(\sum_{e_{\text{in}}:v \text{ input}}\gamma(e_{\text{in}})-\sum_{e_{\text{out}}:v \text{ output}}\gamma(e_{\text{out}})\Big)^2}{\sum_{e\in E}\gamma(e)^2}.\end{aligned}$$

We also define the constant

$$\tau := \max_{e=(v,w):\deg w\geq \deg v}\left(\frac{(\deg w-\deg v+n)\cdot\deg v}{\deg v+\deg w}\right).$$

Theorem 4.2.1. *For every graph,*

$$Q = \max_{\gamma:E\to\mathbb{R}} \frac{\sum_{v\in V}\frac{1}{\deg v}\cdot\Big|\sum_{e_{in}:v \ input}\gamma(e_{in})-\sum_{e_{out}:v \ output}\gamma(e_{out})\Big|}{\sum_{e\in E}|\gamma(e)|}$$

and

$$Q\leq\lambda_n\leq Q\cdot\tau.$$

Observe that the characterization of Q appearing in Theorem 4.2.1 equals the Rayleigh quotient we have used for writing λ_n from the point of view of the edge-Laplacian, replacing the L_2-norm by the L_1-norm. Therefore, such a characterization is analogous to the one of h in Theorem 4.1.1. We prove Theorem 4.2.1 in Section 4.3. Also, in Section 4.4 we motivate the choice of Q, in Section 4.5 we discuss the precision of the lower bound appearing in Theorem 4.2.1 and in Section 4.6 we discuss the precision of the upper bound.

4.3 Proof of the main results

We split the statement of Theorem 4.2.1 into three parts. The first part, Lemma 4.3.1, contains the characterization of Q. The second part, Lemma 4.3.3, states that $Q \leq \lambda_n$. The third part, Lemma 4.3.5, states that $\lambda_n \leq Q \cdot \tau$.

4.3.1 Characterization of Q

Lemma 4.3.1. *For every graph,*

$$Q = \max_{\gamma:E\to\mathbb{R}} \frac{\sum_{v\in V} \frac{1}{\deg v} \cdot \left|\sum_{e_{in}:v \text{ } input} \gamma(e_{in}) - \sum_{e_{out}:v \text{ } output} \gamma(e_{out})\right|}{\sum_{e\in E} |\gamma(e)|}.$$

Proof. In order to prove that

$$Q \leq \max_{\gamma:E\to\mathbb{R}} \frac{\sum_{v\in V} \frac{1}{\deg v} \cdot \left|\sum_{e_{\text{in}}:v \text{ input}} \gamma(e_{\text{in}}) - \sum_{e_{\text{out}}:v \text{ output}} \gamma(e_{\text{out}})\right|}{\sum_{e\in E} |\gamma(e)|},$$

fix an edge (v_1, v_2) that maximizes $\frac{1}{\deg v} + \frac{1}{\deg w}$ over all $(v,w) \in E$ and let $\gamma' : E \to \mathbb{R}$ be 1 on (v_1, v_2) and 0 otherwise. Then,

$$\begin{aligned}
Q &= \frac{1}{\deg v_2} + \frac{1}{\deg v_2} \\
&= \frac{\sum_{v\in V} \frac{1}{\deg v} \cdot \left|\sum_{e_{\text{in}}:v \text{ input}} \gamma'(e_{\text{in}}) - \sum_{e_{\text{out}}:v \text{ output}} \gamma'(e_{\text{out}})\right|}{\sum_{e\in E} |\gamma'(e)|} \\
&\leq \max_{\gamma:E\to\mathbb{R}} \frac{\sum_{v\in V} \frac{1}{\deg v} \cdot \left|\sum_{e_{\text{in}}:v \text{ input}} \gamma(e_{\text{in}}) - \sum_{e_{\text{out}}:v \text{ output}} \gamma(e_{\text{out}})\right|}{\sum_{e\in E} |\gamma(e)|}.
\end{aligned}$$

We now prove that

$$Q \geq \max_{\gamma:E\to\mathbb{R}} \frac{\sum_{v\in V} \frac{1}{\deg v} \cdot \left|\sum_{e_{\text{in}}:v \text{ input}} \gamma(e_{\text{in}}) - \sum_{e_{\text{out}}:v \text{ output}} \gamma(e_{\text{out}})\right|}{\sum_{e\in E} |\gamma(e)|}.$$

Let $\hat{\gamma} : E \to \mathbb{R}$ be a maximizer for

$$\frac{\sum_{v\in V} \frac{1}{\deg v} \cdot \left|\sum_{e_{\text{in}}:v \text{ input}} \gamma(e_{\text{in}}) - \sum_{e_{\text{out}}:v \text{ output}} \gamma(e_{\text{out}})\right|}{\sum_{e\in E} |\gamma(e)|}$$

such that, without loss of generality, $\sum_{e\in E}|\hat\gamma(e)|=1$. Then,

$$\begin{aligned}
Q &= \max_{e=(v,w)}\left(\frac{1}{\deg v}+\frac{1}{\deg w}\right)\\
&= \left(\max_{e=(v,w)}\left(\frac{1}{\deg v}+\frac{1}{\deg w}\right)\right)\cdot\left(\sum_{e\in E}|\hat\gamma(e)|\right)\\
&\geq \sum_{e=(v,w)}|\hat\gamma(e)|\cdot\left(\frac{1}{\deg v}+\frac{1}{\deg w}\right)\\
&= \sum_{v\in V}\frac{1}{\deg v}\cdot\left(\sum_{e:v\text{ input or output}}|\hat\gamma(e)|\right)\\
&\geq \sum_{v\in V}\frac{1}{\deg v}\cdot\left|\sum_{e_{\text{in}}:v\text{ input}}\hat\gamma(e_{\text{in}})-\sum_{e_{\text{out}}:v\text{ output}}\hat\gamma(e_{\text{out}})\right|\\
&= \max_{\gamma:E\to\mathbb{R}}\frac{\sum_{v\in V}\frac{1}{\deg v}\cdot\left|\sum_{e_{\text{in}}:v\text{ input}}\gamma(e_{\text{in}})-\sum_{e_{\text{out}}:v\text{ output}}\gamma(e_{\text{out}})\right|}{\sum_{e\in E}|\gamma(e)|}.
\end{aligned}$$

This proves the claim. □

As a corollary of Lemma 4.3.1, we get another characterization of Q.

Corollary 4.3.2.

$$Q=\max_{\hat\Gamma\subset\Gamma\text{ bipartite}}\frac{\sum_{v\in V}\frac{\deg_{\hat\Gamma}(v)}{\deg_\Gamma(v)}}{|E(\hat\Gamma)|}.$$

Proof. Fix $\Gamma'\subset\Gamma$ that maximizes

$$\frac{\sum_{v\in V}\frac{\deg_{\hat\Gamma}(v)}{\deg_\Gamma(v)}}{|E(\hat\Gamma)|}.$$

over all $\hat\Gamma\subset\Gamma$ bipartite. Fix an orientation and let $\gamma':E(\Gamma)\to\mathbb{R}$ be 1 on each oriented edge in $E(\Gamma')$ and 0 otherwise. Then,

$$\begin{aligned}
Q &= \max_{\gamma:E\to\mathbb{R}}\frac{\sum_{v\in V}\frac{1}{\deg v}\cdot\left|\sum_{e_{\text{in}}:v\text{ input}}\gamma(e_{\text{in}})-\sum_{e_{\text{out}}:v\text{ output}}\gamma(e_{\text{out}})\right|}{\sum_{e\in E}|\gamma(e)|}\\
&\geq \frac{\sum_{v\in V}\frac{1}{\deg v}\cdot\left|\sum_{e_{\text{in}}:v\text{ input}}\gamma'(e_{\text{in}})-\sum_{e_{\text{out}}:v\text{ output}}\gamma'(e_{\text{out}})\right|}{\sum_{e\in E}|\gamma'(e)|}
\end{aligned}$$

$$= \frac{\sum_{v\in V} \frac{\deg_{\Gamma'}(v)}{\deg_\Gamma(v)}}{|E(\Gamma')|}$$

$$= \max_{\hat{\Gamma}\subset\Gamma \text{ bipartite}} \frac{\sum_{v\in V} \frac{\deg_{\hat{\Gamma}}(v)}{\deg_\Gamma(v)}}{|E(\hat{\Gamma})|}.$$

To prove the inverse inequality, let (v_1, v_2) be an edge that maximizes $\frac{1}{\deg v} + \frac{1}{\deg w}$ over all $(v, w) \in E$. Then, by taking $\hat{\Gamma} \subset \Gamma$ as the bipartite graph containing only the edge (v_1, v_2), we get that

$$\max_{\hat{\Gamma}\subset\Gamma \text{ bipartite}} \frac{\sum_{v\in V} \frac{\deg_{\hat{\Gamma}}(v)}{\deg_\Gamma(v)}}{|E(\hat{\Gamma})|} \geq \frac{1}{\deg v_1} + \frac{1}{\deg v_2} = \max_{(v,w)} \left(\frac{1}{\deg v} + \frac{1}{\deg w} \right) = Q.$$

□

4.3.2 Lower bound for the largest eigenvalue

Lemma 4.3.3. *For every graph,*

$$Q \leq \lambda_n.$$

Proof. As in the proof of Lemma 4.3.1, fix an edge (v_1, v_2) that maximizes $\frac{1}{\deg v} + \frac{1}{\deg w}$ over all edges (v, w) and let $\gamma' : E \to \mathbb{R}$ be 1 on (v_1, v_2) and 0 otherwise. Then,

$$\lambda_n = \max_{\gamma: E\to\mathbb{R}} \frac{\sum_{v\in V} \frac{1}{\deg v} \cdot \left(\sum_{e_{\text{in}}: v \text{ input}} \gamma(e_{\text{in}}) - \sum_{e_{\text{out}}: v \text{ output}} \gamma(e_{\text{out}}) \right)^2}{\sum_{e\in E} \gamma(e)^2}$$

$$\geq \frac{\sum_{v\in V} \frac{1}{\deg v} \cdot \left(\sum_{e_{\text{in}}: v \text{ input}} \gamma'(e_{\text{in}}) - \sum_{e_{\text{out}}: v \text{ output}} \gamma'(e_{\text{out}}) \right)^2}{\sum_{e\in E} \gamma'(e)^2}$$

$$= \frac{1}{\deg v_1} + \frac{1}{\deg v_2}$$

$$= Q.$$

□

Remark 4.3.4. *Observe that $Q \geq \frac{n}{n-1}$ if and only if there exists a vertex of degree 1. In fact, if there exists such a vertex, then*

$$Q \geq 1 + \frac{1}{n-1} = \frac{n}{n-1}.$$

If such vertex doesn't exist, then

$$Q \leq \frac{1}{2} + \frac{1}{2} = 1 \leq \frac{n}{n-1}.$$

Therefore, the bound in Lemma 4.3.3 is better than the usual bound $\frac{n}{n-1} \leq \lambda_n$ *only for a small class of graphs. However, the aim of our work is not to find the* best possible bounds for λ_n *but the* best possible bounds for λ_n involving Q*, in order to show that* Q *is a Cheeger-like constant. We shall see, in Section 4.5, that the one in Lemma 4.3.3 is actually the best possible lower bound for* λ_n *involving* Q*.*

4.3.3 Upper bound for the largest eigenvalue

Lemma 4.3.5. *For every graph,*

$$\lambda_n \leq Q \cdot \tau.$$

Proof. Applying [55, Theorem 5], we have that

$$\begin{aligned}
\lambda_n &\leq 2 - \min_{(v,w)} \frac{\left|\mathcal{N}(v) \cap \mathcal{N}(w)\right|}{\max\{\deg v, \deg w\}} \\
&\leq 2 - \min_{(v,w):\deg w \geq \deg v} \frac{\deg v + \deg w - n}{\deg w} \\
&= \max_{(v,w):\deg w \geq \deg v} \frac{\deg w - \deg v + n}{\deg w} \\
&= \max_{(v,w):\deg w \geq \deg v} \left(\frac{1}{\deg v} + \frac{1}{\deg w}\right) \cdot \left(\frac{(\deg w - \deg v + n) \cdot \deg v}{\deg v + \deg w}\right) \\
&\leq Q \cdot \tau.
\end{aligned}$$

□

Observe that the bound in Lemma 4.3.5 is not a better upper bound for λ_n than the one in [55, Theorem 5]. Nevertheless, it is a good upper bound for λ_n involving Q, as we shall see in Section 4.6.

4.4 Choice of Q

We now motivate the choice of Q. As we have discussed in Section 4.1,

$$\lambda_n = \max_{f:V \to \mathbb{R}} \frac{\sum_{v \sim w} \Big(f(v) - f(w)\Big)^2}{\sum_{v \in V} \deg v \cdot f(v)^2} \tag{4.4.1}$$

$$= \max_{\gamma:E\to\mathbb{R}} \frac{\sum_{v\in V} \frac{1}{\deg v} \cdot \left(\sum_{e_{\text{in}}:v \text{ input}} \gamma(e_{\text{in}}) - \sum_{e_{\text{out}}:v \text{ output}} \gamma(e_{\text{out}}) \right)^2}{\sum_{e\in E} \gamma(e)^2}. \tag{4.4.2}$$

We have chosen Q to be the constant that can be written as (4.4.2) by replacing the L_2-norm by the L_1-norm both in the numerator and denominator. We could have chosen to work on the constant that can be written as (4.4.1) by replacing the L_2-norm by the L_1-norm, but such constant is actually equal to 1 for all graphs, as shown by the following lemma. Furthermore, while the characterization of the Cheeger constant is interesting also because it is equal to the second smallest eigenvalue of the 1-*Laplacian*, one cannot get an analogous constant in this sense because the largest eigenvalue of the 1-*Laplacian* equals 1 for every graph, as shown in [16, Theorem 5.1].

Lemma 4.4.1. *For every graph,*

$$\max_{f:V\to\mathbb{R}} \frac{\sum_{v\sim w} \left| f(v) - f(w) \right|}{\sum_{v\in V} \deg v \cdot \left| f(v) \right|} = 1.$$

Proof. Let $\hat{f} : V \to \mathbb{R}$ be a maximizer of

$$\frac{\sum_{v\sim w} \left| f(v) - f(w) \right|}{\sum_{v\in V} \deg v \cdot \left| f(v) \right|}$$

and assume, without loss of generality, that $\sum_{v\in V} \deg v \cdot \left| \hat{f}(v) \right| = 1$. Then,

$$\begin{aligned} \max_{f:V\to\mathbb{R}} \frac{\sum_{v\sim w} \left| f(v) - f(w) \right|}{\sum_{v\in V} \deg v \cdot \left| f(v) \right|} &= \sum_{v\sim w} \left| \hat{f}(v) - \hat{f}(w) \right| \\ &\leq \sum_{v\sim w} \left| \hat{f}(v) \right| + \left| \hat{f}(w) \right| \\ &= \sum_{v\in V} \deg v \cdot \left| \hat{f}(v) \right| \\ &= 1. \end{aligned}$$

To see the inverse inequality, let $\tilde{f} : V \to \mathbb{R}$ that is 1 on a fixed vertex and 0 on all other vertices. Then,

$$\max_{f:V\to\mathbb{R}} \frac{\sum_{v\sim w} \left| f(v) - f(w) \right|}{\sum_{v\in V} \deg v \cdot \left| f(v) \right|} \geq \frac{\sum_{v\sim w} \left| \tilde{f}(v) - \tilde{f}(w) \right|}{\sum_{v\in V} \deg v \cdot \left| \tilde{f}(v) \right|} = 1.$$

□

4.5 How good is the lower bound?

To see that $Q \leq \lambda_n$ is a sharp lower bound, consider the case of K_2: here, $Q = \lambda_2 = 2$. Also, for $n > 2$, consider a non-bipartite graph such that there exists an edge (v, w) with $\deg v = 1$ and $\deg w = 2$. Then, clearly

$$Q = 1 + \frac{1}{2} = \frac{3}{2}$$

and, since the graph is non-bipartite, $\lambda_n < 2$. Therefore, if we look for a bound of the form

$$Q \cdot \nu \leq \lambda_n,$$

we must have

$$\nu \leq \frac{\lambda_n}{Q} < \frac{4}{3} \simeq 1.33.$$

Hence $Q \leq \lambda_n$ is actually a good lower bound involving Q for each n.

4.6 How good is the upper bound?

In order to see that the bound $Q \cdot \tau$ is actually a good upper bound for λ_n, we first construct an example for which the bound $\lambda_n \leq Q \cdot \tau$ is sharp.

Example 4.6.1. *For d-regular graphs, it's easy to see that $Q = \frac{2}{d}$ and $\tau = \frac{n}{2}$, therefore $\lambda_n \leq Q \cdot \tau$ is equivalent to*

$$\lambda_n \leq \frac{n}{d}.$$

In the particular case of the complete graph K_n, $d = n - 1$ and $\lambda_n = \frac{n}{n-1}$ therefore $\lambda_n = Q \cdot \tau$, i.e. the inequality in Lemma 4.3.5 becomes an equality.

For further motivating our upper bound, we shall:

(a) Prove that, for each graph on n nodes,

$$\tau < 0.54 \cdot n$$

and 0.54 is the best ε with a precision of two decimal places such that

$$\lambda_n \leq Q \cdot \varepsilon \cdot n.$$

(b) Prove that we cannot have a bound of the form

$$\lambda_n \leq Q \cdot \left(\frac{n}{2} + c\right),$$

if c is a constant that does not depend on n, as we might be tempted to do by looking at the example of regular graphs.

In order to prove these two points, we shall first discuss *one-sided bipartite graphs*, a new big class of graphs that includes among others petal graphs, complete graphs and complete bipartite graphs.

4.6.1 One-sided bipartite graphs

Definition 4.6.2. *Fix n and k such that $0 < k \leq n-2$. Let $\Gamma = (V, E)$ be a graph on n vertices such that $V = V_1 \sqcup V_2$, $|V_2| = k$ therefore $|V_1| = n - k$, $(v_1, v_2) \in E$ for each $v_1 \in V_1$ and $v_2 \in V_2$, $\deg v_2 = n - k$ for each $v_2 \in V_2$ and $\deg v_1 = d$ for each $v_1 \in V_1$, for some $d \geq k$. Call such a graph a (k, d)-one-sided bipartite graph.*

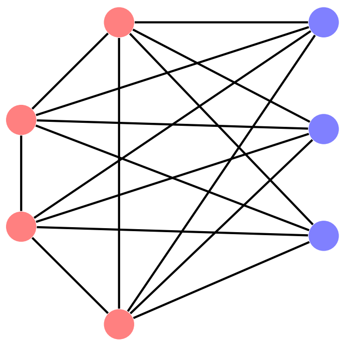


Figure 4.1: A (k, d)-one-sided bipartite graph on 7 nodes, with $k = 3$ and $d = 5$. The red nodes are the ones of degree d.

Remark 4.6.3. *In a (k, d)-one-sided bipartite graph, the vertex set is divided into two sets V_1 and V_2. All possible edges between V_1 and V_2 are there, the k vertices in V_2 are not connected to each other and the vertices in V_1 all have degree d, therefore there are edges between vertices of V_1 if and only if $d > k$ (Figure 4.1). In particular, a (k, d)-one-sided bipartite graph is:*

- *The petal graph if $k = 1$ and $d = 2$;*

- *The complete graph K_n if $k = 1$ and $d = n - 1$;*
- *The complete graph with one edge removed, if $k = 2$ and $d = n - 1$;*
- *The complete bipartite graph $K_{d,n-k}$ if $d = k$;*
- *Not bipartite if $d > k$;*
- *A d-regular graph if $d = n - k$.*

Lemma 4.6.4. *Given n, k and d such that $n \geq 3$, $0 < k \leq n - 2$ and $k \leq d \leq n - 1$, there exists a (k, d)-one-sided bipartite graph on n nodes if and only if at least one of $d - k$ and $n - k$ is even.*

Proof. It follows easily by definition of one-sided bipartite graphs and by [14, Theorem 2.6], that states that a d-regular graph on n nodes exists if and only if at least one of d and n is even. □

In Theorem 4.6.7 we shall prove that for a one-sided bipartite graph with $d \geq n - k$,

$$\lambda_n = \frac{d+k}{d}$$

and for a (k, d)-one-sided bipartite graph with $d < n - k$,

$$\frac{d+k}{d} \leq \lambda_n \leq \frac{n}{d}.$$

We prove a preliminary lemma first.

Definition 4.6.5 ([3])**.** *We say that $v_1, v_2 \in V$ are duplicate vertices if $\mathcal{N}(v_1) = \mathcal{N}(v_2)$.*

Observe that, in particular, duplicate vertices have the same degree and they cannot be neighbors of each other.

Lemma 4.6.6. *If v_1 and v_2 are duplicate vertices and f is an eigenfunction for an eigenvalue $\lambda \neq 1$ of L,*

$$f(v_1) = f(v_2).$$

Proof. Observe that λ eigenvalue of L with eigenfunction f means that, for each vertex v,

$$\lambda \cdot f(v) = Lf(v) = f(v) - \frac{1}{\deg v} \cdot \sum_{v' \sim v} f(v').$$

In particular,

$$\begin{aligned}\lambda \cdot f(v_1) &= f(v_1) - \frac{1}{\deg v_1} \cdot \sum_{v' \sim v_1} f(v') \\ &= f(v_1) - \frac{1}{\deg v_2} \cdot \sum_{v' \sim v_2} f(v')\end{aligned}$$

and

$$\lambda \cdot f(v_2) = f(v_2) - \frac{1}{\deg v_2} \cdot \sum_{v' \sim v_2} f(v').$$

Therefore,

$$\frac{1}{\deg v_2} \cdot \sum_{v' \sim v_2} f(v') = f(v_1) \cdot (1-\lambda) = f(v_2) \cdot (1-\lambda).$$

Since by assumption $\lambda \neq 1$, this implies that $f(v_1) = f(v_2)$. □

Theorem 4.6.7. *For a (k,d)-one-sided bipartite graph with $d \geq n-k$,*

$$\lambda_n = \frac{d+k}{d}.$$

For a (k,d)-one-sided bipartite graph with $d < n-k$,

$$\frac{d+k}{d} \leq \lambda_n \leq \frac{n}{d}.$$

Proof. For any fixed (k,d)-one-sided bipartite graph, let $\lambda \neq 0, 1$ be an eigenvalue for L with eigenfunction f. By construction, in a (k,d)-one-sided bipartite graph all k vertices in V_2 of degree $n-k$ are duplicate vertices. Therefore, by Lemma 4.6.6, $f(v_2)$ is constant for each $v_2 \in V_2$. If, in particular, $f(v_2) \neq 0$ for each $v_2 \in V_2$, we can define

$$\alpha_{v_2} := \frac{-\sum_{v_1 \in V_1} f(v_1)}{f(v_2)}$$

and, since this is constant for each $v_2 \in V_2$, we can write $\alpha_{n-k} = \alpha_{v_2}$. Therefore,

$$\lambda \cdot f(v_2) = f(v_2) - \frac{1}{n-k} \cdot \sum_{v_1 \in V_1} f(v_1) = f(v_2) \cdot \left(1 + \frac{\alpha_{n-k}}{n-k}\right),$$

which implies that

$$\lambda = 1 + \frac{\alpha_{n-k}}{n-k}.$$

In particular, since we are assuming $\lambda \neq 1$, this implies that $\alpha_{n-k} \neq 0$, hence we can write

$$f(v_2) = \frac{-\sum_{v_1 \in V_1} f(v_1)}{\alpha_{n-k}}.$$

Now, by the orthogonality to the constants, we must have $\sum_v \deg v \cdot f(v) = 0$. Hence

$$0 = \sum_{v_1 \in V_1} d \cdot f(v_1) + k \cdot (n-k) \cdot \left(\frac{-\sum_{v_1 \in V_1} f(v_1)}{\alpha_{n-k}} \right)$$
$$= \left(\sum_{v_1 \in V_1} f(v_1) \right) \cdot \left(d - \frac{k \cdot (n-k)}{\alpha_{n-k}} \right).$$

If

$$\sum_{v_1 \in V_1} f(v_1) = 0,$$

then $\alpha_{n-k} = 0$ therefore $\lambda = 1$, which is a contradiction. Therefore we must have

$$d - \frac{k \cdot (n-k)}{\alpha_{n-k}} = 0,$$

which implies that

$$\alpha_{n-k} = \frac{k \cdot (n-k)}{d}$$

therefore

$$\lambda = 1 + \frac{k}{d} = \frac{d+k}{d}.$$

This proves that $\frac{d+k}{d}$ is an eigenvalue, therefore

$$\lambda_n \geq \frac{d+k}{d}.$$

Now, in the particular case of $d \geq n-k$, we can prove also the inverse inequality by applying [55, Theorem 5], that states that

$$\lambda_n \leq 2 - \min_{(v,w)} \frac{\left|\mathcal{N}(v) \cap \mathcal{N}(w)\right|}{\max\{\deg v, \deg w\}}.$$

We prove that, for a (k, d)-one-sided bipartite graph with $d \geq n-k$,

$$\min_{(v,w)} \frac{\left|\mathcal{N}(v) \cap \mathcal{N}(w)\right|}{\max\{\deg v, \deg w\}} = \frac{d-k}{d}.$$

We consider the possible cases.

- Case 1: $v \in V_1$ and $w \in V_2$. Since we are assuming $d \geq n-k$, we have that $\max\{\deg v, \deg w\} = d$. Therefore,

$$\frac{\left|\mathcal{N}(v) \cap \mathcal{N}(w)\right|}{\max\{\deg v, \deg w\}} = \frac{d-k}{d}.$$

- Case 2: $v, w \in V_1$. In this case, $\deg v = \deg w = d$. Also, v and w have k neighbors in common in V_2 and at least $2(d-k)-(n-k)$ neighbors in common in V_1. Therefore,

$$\frac{\left|\mathcal{N}(v) \cap \mathcal{N}(w)\right|}{\max\{\deg v, \deg w\}} \geq \frac{k+2(d-k)-(n-k)}{d} = \frac{2d-n}{d} \geq \frac{d-k}{d},$$

where the last inequality follows from the assumption that $d \geq n-k$.

Therefore,

$$\min_{(v,w)} \frac{\left|\mathcal{N}(v) \cap \mathcal{N}(w)\right|}{\max\{\deg v, \deg w\}} = \frac{d-k}{d}$$

and by [55, Theorem 5] this implies that

$$\lambda_n \leq 2 - \frac{d-k}{d} = \frac{d+k}{d},$$

therefore that the equality holds in this case.

It remains to prove that, for $d < n-k$,

$$\lambda_n \leq \frac{n}{d}.$$

Let again $\lambda \neq 0, 1$ be an eigenvalue for L with eigenfunction f. We know that $f(v_2)$ must be constant for each $v_2 \in V_2$, and we have already seen the case $f(v_2) \neq 0$. Consider now the case $f(v_2) = 0$. We have that

$$\begin{aligned}
\lambda &= \frac{\sum_{v \sim w} \Big(f(v) - f(w)\Big)^2}{\sum_{v_1 \in V} d \cdot f(v_1)^2} \\
&= \frac{\sum_{v_1 \in V_1} k \cdot f(v_1)^2 + \sum_{v \sim w \in V_1} \Big(f(v) - f(w)\Big)^2}{\sum_{v_1 \in V} d \cdot f(v_1)^2} \\
&= \frac{k}{d} + \frac{\sum_{v \sim w \in V_1} \Big(f(v) - f(w)\Big)^2}{\sum_{v_1 \in V} d \cdot f(v_1)^2} \\
&\leq \frac{k}{d} + \lambda_n',
\end{aligned}$$

where λ_n' is the largest eigenvalue of a d-regular graph on $n-k$ nodes, therefore

$$\lambda_n' \leq \frac{n-k}{d}.$$

In fact, in order to prove it it's enough to show that, for d-regular graphs on $\hat{n}$ nodes, the largest eigenvalue of the non-normalized Laplace operator is at most $\hat{n}$. This is actually true for every graph, because for the non-normalized Laplacian the complete graph has largest eigenvalue equal to $\hat{n}$ and, *if an edge is added into a graph, then none of its Laplacian eigenvalues can decrease* [42]. Therefore,

$$\lambda \leq \lambda_n \leq \frac{k}{d} + \frac{n-k}{d} = \frac{n}{d}$$

$\square$

Remark 4.6.8. *Observe also that, for (k, d)-one-sided bipartite graphs with $d \geq n - k$,*

$$Q = \frac{1}{d} + \frac{1}{n-k}.$$

For (k, d)-one-sided bipartite graphs with $d < n - k$,

$$Q = \frac{2}{d}.$$

4.6.2 Conclusions

As a consequence of Theorem 4.6.7, we can prove the following corollary that further motivates the upper bound in Lemma 4.3.5.

Corollary 4.6.9. *(a) For each graph on n nodes,*

$$\tau < 0.54 \cdot n$$

and 0.54 *is the best ε with a precision of two decimal places such that*

$$\lambda_n \leq Q \cdot \varepsilon \cdot n.$$

(b) We cannot have a bound of the form

$$\lambda_n \leq Q \cdot \left(\frac{n}{2} + c\right),$$

if c is a constant that does not depend on n.

Proof. (a) By writing in WolframAlpha [61]:

```
(y(z-y+x))/(y+z) >= 0.54 x
with x>0, y>0, y<x, z>=y, z<x, integer solutions
```

one can see that there is no solution. Therefore,

$$\tau < 0.54 \cdot n$$

for each graph and, by Lemma 4.3.5, $\lambda_n \leq Q \cdot 0.54 \cdot n$. In order to see that 0.54 is the best ε with a precision of two decimal places such that

$$\lambda_n \leq Q \cdot \varepsilon \cdot n,$$

observe that for (k,d)-one-sided bipartite graphs with $d \geq n-k$, we have that

$$\frac{\lambda_n}{Q} = \frac{dn - dk + kn - k^2}{d + n - k}.$$

By writing in WolframAlpha [61]:

```
(xz-yz+xy-y^2)/(x-y+z) > (0.53*x),
with x>0, y>0, y<x-1, z>=x-y, z>=y, z<x integer solutions
```

one can see that there are solutions, for example for $x = n = 249$, $y = k = 69$ and $z = d = 241$. Since $n-k$ is even, by Lemma 4.6.4 there exists a (k,d)-one-sided bipartite graph with these values of n, k and d. For such a graph,

$$\lambda_n > Q \cdot 0.53 \cdot n.$$

This proves the first claim.

(b) For (k,d)-one-sided bipartite graphs with $d = n-1$ and $k = \frac{n}{4}$,

$$\frac{\lambda_n}{Q} = \frac{15n^2 - 12n}{28n - 16}.$$

Therefore, if we look for an upper bound of λ_n of the form $Q \cdot g(n)$, we must have $g(n) \geq \frac{15n^2-12n}{28n-16}$ for each n. In particular, we cannot take any $g(n) = \frac{n}{2} + c$ if c is a constant that does not depend on n.

$\square$

5 Hypergraph Laplace operators for chemical reaction networks

In this chapter we define *chemical hypergraphs* as a new model for representing chemical reaction networks. We also define two Laplace operators for chemical hypergraphs that generalize L and L^V and we investigate some properties of their spectra. These results are presented also in [38], a joint work with Jürgen Jost.

5.1 First definitions and assumptions

Chemical reaction networks can be modelled by *directed* hypergraphs, in which each vertex represents a chemical element and each hyperedge represents a chemical reaction involving the elements that it contains as vertices. Each reaction is a directed hyperedge, mapping a collection of vertices, its educts or inputs, to another collection, its products or outputs. We could therefore define a suitable Laplace type operator for a directed hypergraph and study its spectrum, as pioneered by Frank Bauer [6] for directed graphs. Since such an operator is not self-adjoint with respect to some scalar product, however, in general its eigenvalues will not be real, but have non-zero imaginary parts. Here, however, we prefer to work with symmetric operators and real eigenvalues. That would suggest to work with undirected hypergraphs. Nevertheless, we preserve an important bit of additional structure from the chemical reaction networks, the fact that the vertex set of a hyperedge is partitioned into two classes. In the directed case, they correspond to inputs and outputs, but in the setting that we wish to adopt, we do not distinguish these two roles and simply keep the partitioning of the vertices of a hyperedge into two classes. Thus, we are working with hypergraphs with an additional piece of structure, the partitioning of the vertex sets of each hyperedge into two classes. We shall call these *chemical hypergraphs.*

Definition 5.1.1. *A chemical hypergraph is a pair* $\Gamma = (V, H)$ *such that* $V = \{v_1, \ldots, v_N\}$ *is a finite set of vertices and* H *is a set such that every element* h

in H is a pair of elements (V_h, W_h) (input and output, not necessarily disjoint) in $\mathcal{P}(V) \setminus \{\emptyset\}$. The elements of H are called the oriented hyperedges. Changing the orientation of a hyperedge h means exchanging its input and output, leading to the pair (W_h, V_h).

From now on, we shall simply call them hypergraphs.
We assume that each vertex is contained in at least one hyperedge. Also, since every chemical reaction has both educts and products, we consider only hyperedges that have at least one input and at least one output.

Definition 5.1.2. *A catalyst in a hyperedge h is a vertex that is both an input and an output for h (Figure 5.1).*

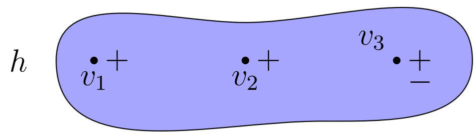


Figure 5.1: An hyperedge h that has two inputs and one catalyst.

Remark 5.1.3. *The above definition comes from the fact that, in chemistry, a catalyst is an element that participates in a reaction but is not changed by that reaction.*

Our theory thus includes also *graphs with self-loops*, i.e. graphs that may have edges whose two endpoints coincide.

Analogously to the case of graphs, while we shall not work with *directed* hyperedges, we shall nevertheless have to work with *oriented* hyperedges. That is, when h is a hyperedge with its two vertex sets V_h, W_h, it can carry two orientations, one going from V_h to W_h and the other in the opposite direction, from W_h to V_h. We arbitrarily call the two orientations of h $+$ and $-$. We shall consider functions γ from the set of oriented hyperedges that satisfy

$$\gamma(h, -) = -\gamma(h, +). \tag{5.1.1}$$

Definition 5.1.4. *We say that two hypergraphs $\Gamma = (V, H)$ and $\Gamma = (V', H')$ are isomorphic if there exist two bijections,*

$$\varphi_V : V \to V' \qquad \text{and} \qquad \varphi_H : H \to H'$$

such that, for each hyperedge $h \in H$,

$$h = (V_h, W_h) \qquad \Longleftrightarrow \qquad \varphi_H(h) = (\varphi_V(V_h), \varphi_V(W_h)).$$

Definition 5.1.5. *We say that a hypergraph* $\Gamma = (V, H)$ *is connected (Figure 5.2) if, for every pair of vertices* $v, w \in V$, *there exists a path that connects* v *and* w, *i.e. there exist* $v_1, \ldots, v_m \in V$ *and* $h_1, \ldots, h_{m-1} \in H$ *such that:*

- $v_1 = v$;
- $v_m = w$;
- $\{v_i, v_{i+1}\} \subseteq h_i$ *for each* $i = 1, \ldots, m-1$.

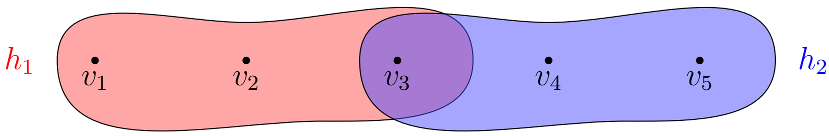


Figure 5.2: A connected hypergraph.

Definition 5.1.6. *We say that* $\Gamma = (V, H)$ *has* k *connected components if there exist* $\Gamma_1 = (V_1, H_1), \ldots, \Gamma_k = (V_k, H_k)$ *such that:*

(a) For every $i \in \{1, \ldots, k\}$, Γ_i *is a connected hypergraph with* $V_i \subseteq V$ *and* $H_i \subseteq H$;

(b) For every $i, j \in \{1, \ldots, k\}$, $i \neq j$, $V_i \cap V_j = \emptyset$ *and therefore also* $H_i \cap H_j = \emptyset$.

(c) $\bigcup V_i = V$, $\bigcup H_i = H$.

Definition 5.1.7. *Let* $\Gamma = (V, H)$ *be a hypergraph. We say that* $\mathcal{S} = (V', H')$ *is a closed system of reactions in* Γ *(Figure 5.3) if:*

(a) $\emptyset \neq H' \subseteq H$;

(b) $V' = \{v \in h : h \in H'\}$;

(c) Each $v \in V'$ *appears in* $\mathcal{S}$ *as often as input as as output.*

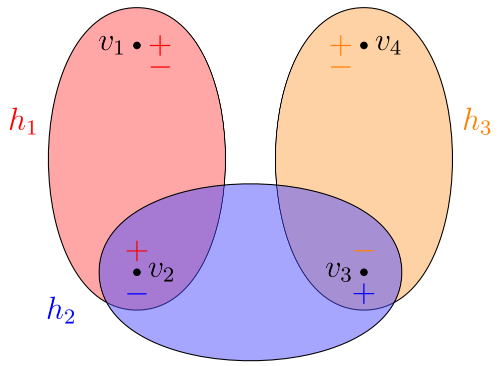


Figure 5.3: A closed system of reactions.

Remark 5.1.8. *Closed systems for hypergraphs generalize the oriented cycles that we have for graphs, so they are interesting from the mathematical point of view, and they are also clearly interesting from the chemical point of view.*

Definition 5.1.9. *We say that two closed systems $\mathcal{S} = (V', H')$ and $\mathcal{S} = (V', H')$ are disjoint if $H \cap H' = \emptyset$.*

Remark 5.1.10. *Disjoint systems don't have common hyperedges but they may have common vertices.*

Definition 5.1.11. *Let $\Gamma = (V, H)$ be a hypergraph with M hyperedges $h_1, \dots, h_M$ and K closed systems of reactions $\mathcal{S}_1, \dots, \mathcal{S}_K$. Let $A = (a_{ij})_{ij}$ be the $K \times M$ matrix such that*

$$a_{ij} := \begin{cases} 1 & \text{if } h_j \in \mathcal{S}_i \\ 0 & \text{otherwise.} \end{cases}$$

Therefore each row A_i of A represents a closed system $\mathcal{S}_i$ and each column A^j of A represents a hyperedge h_j. Given $I \subseteq \{1, \dots, K\}$, we say that the closed systems $\{\mathcal{S}_i\}_{i \in I}$ are linearly independent if the rows $\{A_i\}_{i \in I}$ of A are linearly independent.

Remark 5.1.12. *Pairwise disjoint closed systems are linearly independent.*

5.2 Generalized Laplace Operators

In order to define the Laplace operators for hypergraphs, we will generalize the construction of the Laplace operators for graphs presented in Chapter 2 in the most natural way. In particular, we will:

(i) Give weight one to the hyperedges (as we do for edges in the case of graphs) and therefore give weight $\deg v := \left|\text{hyperedges containing } v\right|$ to each vertex v;

(ii) Define a scalar product for functions defined on hyperedges and a scalar product for functions defined on vertices, based on the weights we gave;

(iii) Define the boundary operator for functions defined on the vertex set;

(iv) Find the coboundary operator based on the scalar products we defined;

(v) Define the Laplace operators as the two different compositions of the boundary and the coboundary operator.

Definition 5.2.1 (Scalar product for functions defined on oriented hyperedges)**.** *Given $\omega, \gamma : H \to \mathbb{R}$, let*

$$(\omega, \gamma)_H := \sum_{h \in H} \omega(h) \cdot \gamma(h).$$

Definition 5.2.2 (Scalar product for functions defined on vertices)**.** *Given $f, g : V \to \mathbb{R}$, let*

$$(f, g)_V := \sum_{v \in V} \deg v \cdot f(v) \cdot g(v).$$

Definition 5.2.3 (Boundary operator for functions defined on vertices)**.** *Given $f : V \to \mathbb{R}$ and $h \in H$, let*

$$\delta f(h) := \sum_{v_i \text{ input of } h} f(v_i) - \sum_{v^j \text{ output of } h} f(v^j).$$

Remark 5.2.4. *Note that*

$$\delta : C(V) \longrightarrow C(H)$$

where the functions in $C(H)$ are always supposed to satisfy (5.1.1)*. In particular, δf also satisfies* (5.1.1)*.*

Definition 5.2.5 (Adjoint operator of the boundary operator)**.** *Let*

$$\delta^* : C(H) \longrightarrow C(V)$$

be defined as

$$\delta^*(\gamma)(v) := \frac{\sum_{h_{in}: v \text{ input}} \gamma(h_{in}) - \sum_{h_{out}: v \text{ output}} \gamma(h_{out})}{\deg v}.$$

Lemma 5.2.6. *δ^* satisfies $(\delta f, \gamma)_H = (f, \delta^* \gamma)_V$, therefore it is the (unique) adjoint operator of δ.*

Proof.

$$\begin{aligned}
(\delta f, \gamma)_H &= \sum_{h \in H} \gamma(h) \cdot \left(\sum_{v_i \text{ input of } h} f(v_i) - \sum_{v^j \text{ output of } h} f(v^j) \right) \\
&= \sum_{v \in V} f(v) \cdot \left(\sum_{h_{\text{in}}: v \text{ input}} \gamma(h_{\text{in}}) - \sum_{h_{\text{out}}: v \text{ output}} \gamma(h_{\text{out}}) \right) \\
&= \sum_{v \in V} \deg v \cdot f(v) \cdot \frac{\left(\sum_{h_{\text{in}}: v \text{ input}} \gamma(h_{\text{in}}) - \sum_{h_{\text{out}}: v \text{ output}} \gamma(h_{\text{out}}) \right)}{\deg v} \\
&= \sum_{v \in V} \deg v \cdot f(v) \cdot \delta^*(\gamma)(v) \\
&= (f, \delta^* \gamma)_V.
\end{aligned}$$

□

Definition 5.2.7 (Laplace operators)**.** *Given $f : V \to \mathbb{R}$ and given $v \in V$, let*

$$\begin{aligned}
L^V f(v) :=& \delta^*(\delta f)(v) \\
=& \frac{\sum_{h_{in}: v \ input} \delta f(h_{in}) - \sum_{h_{out}: v \ output} \delta f(h_{out})}{\deg v} \\
=& \frac{\sum_{h_{in}: v \ input} \left(\sum_{v' \ input \ of \ h_{in}} f(v') - \sum_{w' \ output \ of \ h_{in}} f(w') \right)}{\deg v} + \\
& - \frac{\sum_{h_{out}: v \ output} \left(\sum_{\hat{v} \ input \ of \ h_{out}} f(\hat{v}) - \sum_{\hat{w} \ output \ of \ h_{out}} f(\hat{w}) \right)}{\deg v}.
\end{aligned}$$

Analogously, given $\gamma : H \to \mathbb{R}$ and $h \in H$, let

$$\begin{aligned}
L^H \gamma(h) :=& \delta(\delta^* \gamma)(h) \\
=& \sum_{v_i \ input \ of \ h} \delta^* \gamma(v_i) - \sum_{v^j \ output \ of \ h} \delta^* \gamma(v^j) \\
=& \sum_{v_i \ input \ of \ h} \frac{\sum_{h_{in}: v_i \ input} \gamma(h_{in}) - \sum_{h_{out}: v_i \ output} \gamma(h_{out})}{\deg v_i} +
\end{aligned}$$

$$- \sum_{v^j \text{ output of } h} \frac{\sum_{h'_{in}:v^j \text{ input}} \gamma(h'_{in}) - \sum_{h'_{out}:v^j \text{ output}} \gamma(h'_{out})}{\deg v^j}.$$

Proposition 5.2.8. *For every graph,*

$$L^V = L \qquad \text{and} \qquad L^H = L^E.$$

Proof. It follows from the fact that the restriction of the boundary and coboundary operators for hypergraphs clearly coincide with the boundary and coboundary operators for graphs. □

Remark 5.2.9. *$L^H\gamma(h)$ counts what flows out at the inputs $-$ what flows in at the inputs $-$ what flows out at the outputs $+$ what flows in at the outputs.*

5.3 First properties

Analogously to the case of graphs, we can prove the following lemmas.

Lemma 5.3.1. *L^V and L^H are both self-adjoint.*

Proof. Use the fact that L^V and L^H are the two compositions of δ and δ^*, which are adjoint to each other. □

Lemma 5.3.2. *L^V and L^H are non-negative operators.*

Proof. Let $f : V \to \mathbb{R}$. Then

$$(L^V f, f)_V = (\delta^*\delta f, f)_V = (\delta f, \delta f)_H \geq 0. \tag{5.3.1}$$

Analogously, for $\gamma : H \to \mathbb{R}$,

$$(L^H\gamma, \gamma)_H = (\delta\delta^*\gamma, \gamma)_H = (\delta^*\gamma, \delta^*\gamma)_V \geq 0. \tag{5.3.2}$$

□

A direct consequence of Lemmas 5.3.1 and 5.3.2 is

Corollary 5.3.3. *The eigenvalues of L^V and L^H are real and non-negative.*

Notation 5.3.4. *Let $N := |V|$ and let $M := |H|$. Since the space of real functions on a set with cardinality k is k-dimensional, a self-adjoint operator on* *this space has precisely k eigenvalues, counted with their multiplicities. Therefore L^V has N eigenvalues that we will arrange as*

$$\lambda_1 \leq \ldots \leq \lambda_N$$

Analogously, L^H has M eigenvalues that we will arrange as

$$\lambda_1^H \leq \ldots \leq \lambda_M^H.$$

Again as in the case of graphs, as a corollary of Lemma 2.2.1 we get the following.

Corollary 5.3.5. *The non-zero eigenvalues of L^V and L^H are the same. In particular, if f is an eigenfunction of L^V with eigenvalue $\lambda \neq 0$, then δf is an eigenfunction of L^H with eigenvalue λ; if γ is an eigenfunction of L^H with eigenvalue $\lambda' \neq 0$, then $\delta^* \gamma$ is an eigenfunction of L^V with eigenvalue λ'.*

Therefore, we have two alternative ways to control or estimate the non-vanishing eigenvalues. Furthermore, as in the case of graphs, the two operators only differ in the multiplicity of the eigenvalue 0. Let m_V and m_H be the multiplicity of the eigenvalue 0 of L^V and L^H, respectively. Then Corollary 5.3.5 implies

Corollary 5.3.6.

$$m_V - m_H = |V| - |H|. \tag{5.3.3}$$

In particular, (5.3.3) offers a generalization of the Euler characteristic to hypergraphs. Although we don't see hypergraphs as topological objects due to the irregularities of their structure, (5.3.3) tells us that we can anyway define the Euler characteristic, an important invariant for topological spaces, and the spectra of the Laplace operators capture it.

5.4 The eigenvalue 0

In this section, we want to control the multiplicity of the eigenvalue 0 for our two Laplacians. They are related by Corollary 5.3.6. In order to see the principle, start with the simple situation where we only have a set V of vertices, but no (hyper)edges connecting them. Then (5.3.3) tells us that $m_V = |V|$, which of course can be trivially verified. Now we add edges. As long as these edges do not form cycles, that is, as long as the graph is a forest, i.e., a collection of trees, we have $m_H = 0$, and therefore, each new edge reduces the number of components as well as $m_V = |V| - |H|$ by 1. When, however, a new edge closes a cycle, then m_H increases by 1, and consequently, m_V is left unchanged. A special case of this is when we add a loop to a vertex. A loop induces a new eigenvalue 0 of L^H and thus lets m_V unchanged. The general formula says that $m_V - m_H$ equals the number of connected components minus the number of independent cycles, including self-loops.

Something analogous happens when we more generally add hyperedges. In contrast to the case of graphs, however, by adding hyperedges, we can potentially eliminate all eigenvalues 0 of L^V. For a graph, L^V always has the eigenvalue 0, as should be clear from the preceding or also follows from Lemma 5.4.6 below. We shall see examples of hypergraphs where L^V has only positive eigenvalues. But we first make some obvious observations.

Lemma 5.4.1. *On a hypergraph with a single hyperedge, L^V has 0 as an eigenvalue. More precisely, $m_V = |V|-1$ if not every vertex is a catalyst and $m_V = |V|$ if every vertex is a catalyst.*

Proof. In Example 5.5.1 we shall see that, on a hypergraph with a single hyperedge, the only eigenvalue of L^H is non-zero if and only if not every vertex is a catalyst. Therefore, by (5.3.3), $m_V = |V|-1$ if not every vertex is a catalyst and $m_V = |V|$ otherwise. □

In order to investigate this in more detail, we observe that by (5.3.1), a function f on the vertex set satisfies $L^V f = 0$ if and only for every $h \in H$,

$$\sum_{v_i \text{ input of } h} f(v_i) = \sum_{v^j \text{ output of } h} f(v^j). \tag{5.4.1}$$

Thus, to create an eigenvalue 0 of L^V, we need a function $f : V \to \mathbb{R}$ that is not identically 0 and satisfies (5.4.1).

Similarly, by (5.3.2), in order to get an eigenvalue 0 of L^H, we need $\gamma : H \to \mathbb{R}$ satisfying (5.1.1) and

$$\sum_{h_{\text{in}}:i \text{ input}} \gamma(h_{\text{in}}) = \sum_{h_{\text{out}}:i \text{ output}} \gamma(h_{\text{out}}) \tag{5.4.2}$$

for every vertex i.
And the multiplicity of the eigenvalue 0 of L^V and L^H then is given by the number of linearly independent, or since we have scalar products, equivalently

by the number of orthogonal f and γ, respectively, satisfying these equations. Conversely, if there is no such f or γ, then the corresponding multiplicity is 0.

For instance, (5.4.1) already implies

Lemma 5.4.2. *If a hypergraph has a vertex v_0 that is a catalyst for every hyperedge that it is contained in, then L^V has 0 as an eigenvalue.*

Proof. Let $f(v_0) = 1$ and $f(v) = 0$ for $v \neq v_0$. This then satisfies (5.4.1). □

Remark 5.4.3. *Any function $f : V \to \mathbb{R}$ is an eigenfunction for the eigenvalue 0 in some hypergraph that has vertex set V. Construct a hypergraph Γ in which all the vertices $v_1, \ldots, v_k$ of V are always catalysts. Then f satisfies (5.4.1) for Γ.*

In fact, we have

Proposition 5.4.4. *If k vertices are always catalysts, then $m_V \geq k$.*
And $m_V = N$, or equivalently, $\lambda_N = 0$, that is, 0 is the only eigenvalue $\iff$ all vertices are always catalysts.

Proof. The first part and the implication $\Longleftarrow$ are clear from the proof of Lemma 5.4.2. In order to prove $\Longrightarrow$, we assume that there exists at least one vertex $\hat{v} \in \hat{h}$ which is not a catalyst for $\hat{h}$ (without loss of generality, we can assume that it is an input). We want to prove that $\lambda_N > 0$. Let $f : V \to \mathbb{R}$ such that $f(w) = 0$ for all $w \neq \hat{v}$ and such that

$$f(\hat{v}) = \frac{1}{\sqrt{\deg \hat{v}}}.$$

Then $\sum_{v \in V} \deg v \cdot f(v)^2 = 1$ and

$$\begin{aligned}
&\sum_{h \in H} \Bigg(\sum_{v_i \text{ input of } h} f(v_i) - \sum_{v^j \text{ output of } h} f(v^j) \Bigg)^2 \\
&\geq \Bigg(\sum_{v_i \text{ input of } \hat{h}} f(v_i) - \sum_{v^j \text{ output of } \hat{h}} f(v^j) \Bigg)^2 \\
&= \left(\frac{1}{\sqrt{\deg \hat{v}}} \right)^2 \\
&> 0.
\end{aligned}$$

Therefore, $\lambda_N > 0$. □

Remark 5.4.5. *The previous proposition implies that, unlike the case of the graphs, the multiplicity of the eigenvalue* 0 *for* L^V *is in general not equal to the number of connected components of the hypergraph (in particular, we don't have that* $\lambda_2 > 0$ *for every connected hypergraph) and, analogously, the multiplicity of the eigenvalue* 0 *for* L^H *does not count, in general, the cycles of the hypergraph.*

We shall now see some further special cases of hypergraphs with $\lambda_1 = 0$.

Lemma 5.4.6. *Let* Γ *satisfy*

$$\left|\textit{inputs of } h\right| = \left|\textit{outputs of } h\right| \textit{ for each } h \in H. \tag{5.4.3}$$

Then L^V *has the eigenvalue* 0*.*

This holds in particular for graphs, because there, every edge has precisely one input and one output.

Proof. When (5.4.3) holds, then any constant function satisfies (5.4.1). □

Remark 5.4.7. *In fact, some such condition is necessary. More precisely, the fact that* $\lambda_1 = 0$ *for a hypergraph means that we can give a* weight $f : V \to \mathbb{R}$ *to the vertices such that, in each hyperedge, inputs and outputs have in total the same weight.*

Proposition 5.4.8. *If* Γ *is one of the following hypergraphs, then* $\lambda_1 = 0$*:*

(i) Γ *is given by the union of a hypergraph* Γ' *with* $\lambda_1' = 0$ *together with a hyperedge* h *such that there exists at least one* $v \in h \setminus \Gamma'$*;*

(ii) Γ *is given by the union of a hypergraph* Γ' *with* $\lambda_1' = 0$ *together with a hyperedge* h *that involves only vertices of* Γ' *and has only catalysts.*

Proof. (i) Assume that Γ is given by the union of a hypergraph Γ' with $\lambda_1' = 0$ together with a hyperedge h which involves at least one vertex that is not in Γ'. Since $\lambda_1' = 0$, there exists a function f' for Γ' that satisfies (5.4.1). If there is a vertex in $h \setminus \Gamma'$ which is a catalyst, we can apply Lemma 5.4.2. If h involves at least one vertex $\hat{v} \notin \Gamma'$ which is not a catalyst, let

$$f(v) := f'(v)$$

for every $v \in \Gamma'$;

$$f(w) := 0$$

for every vertex $w \in h \setminus \Gamma'$, $w \neq \hat{v}$;

$$f(\hat{v}) := \sum_{v^j \in \Gamma' : v^j \text{ output of } h} f'(v^j) - \sum_{v_i \in \Gamma' : v_i \text{ input of } h} f'(v_i)$$

if $\hat{v}$ is an input and not an output;

$$f(\hat{v}) := \sum_{v_i \in \Gamma' : v_i \text{ input of } h} f'(v_i) - \sum_{v^j \in \Gamma' : v^j \text{ output of } h} f'(v^j)$$

if $\hat{v}$ is an output and not an input.
Then f satisfies (5.4.1).

(ii) Assume that Γ is given by the union of a hypergraph Γ' with $\lambda_1' = 0$ together with a hyperedge h which involves only vertices of Γ' and which has only catalysts. Since $\lambda_1' = 0$, there exists a function f' for Γ' that satisfies (5.4.1). Such f' satisfies (5.4.1) also for Γ.

□

We shall now see two examples of hypergraphs with $\lambda_1 > 0$, that is, where L^V does not have 0 as an eigenvalue.

Lemma 5.4.9. *Let Γ be the union of a connected graph Γ' with a hyperedge h that involves only vertices of Γ' and such that $\left|\text{inputs of } h\right| \neq \left|\text{outputs of } h\right|$. Then $\lambda_1 > 0$.*

Proof. We know that f satisfies (5.4.1) on a connected graph Γ' if and only if f is a constant function. But a constant function f can clearly not satisfy (5.4.1) for a hyperedge h such that $\left|\text{inputs of } h\right| \neq \left|\text{outputs of } h\right|$. Therefore, λ_1 cannot be 0 in this case. □

Lemma 5.4.10. *Let Γ be the hypergraph on $N > 2$ vertices $v_1, \ldots, v_N$ with N hyperedges $h_1, \ldots, h_N$ such that, for each $i \in \{1, \ldots, N\}$, h_i has:*

- *v_i as input, and*
- *every v_j with $j \neq i$ as output.*

Then $\lambda_1 > 0$.

Proof. Let $f : V \to \mathbb{R}$ be a function that satisfies (5.4.1). Then for every $i, l \in \{1, \ldots, N\}$,

$$f(v_i) = \sum_{j \neq i} f(v_j) = f(v_l) + \sum_{j \neq i,l} f(v_j) = f(v_i) + 2 \cdot \sum_{j \neq i,l} f(v_j).$$

Therefore $\sum_{j\neq i,l} f(v_j) = 0$ and $f(v_i) = f(v_l)$. Since this is true for every $i, l \in \{1, \ldots, N\}$, f must be the zero function. This implies that $\lambda_1 > 0$. $\square$

We now see how to apply (5.4.2). First, when we have a closed system of reactions, we can take a γ that has the same non-zero value on all hyperedges involved in that system and vanishes on all other hyperedges. Such a γ then satisfies (5.4.2) because every vertex in such a system appears the same number of times as input as as output for hyperedges belonging to that system. This is formalized in the next Lemma.

Lemma 5.4.11. *If Γ has a closed system of reactions, then $\lambda_1^H = 0$.*

Proof. Let $\mathcal{S} = (V', H')$ be a closed system in Γ. Let $\gamma : H \to \mathbb{R}$ be defined as $\gamma(h') := 1$ for all $h' \in H'$ and $\gamma(h) := 0$ for all $h \in H \backslash H'$. Then γ satisfies (5.4.2). Therefore $\lambda_1^H = 0$. $\square$

Remark 5.4.12. *The claim of Lemma 5.4.11 is actually an* if and only if *for both the case of graphs (for which we know that the multiplicity of* 0 *for L^H is equal to the number of oriented cycles) and the case of Γ containing only a single hyperedge. In fact, as we shall see in Example 5.5.1, in this case $\lambda_1^H = 0$ if and only if all vertices are catalysts, that is, if and only if there is a closed system of reactions in Γ (which is Γ itself). But Example 5.4.14 will show that the converse of Lemma 5.4.11 does not hold.*

In order to prepare that example, we shall first present another example of a closed system of reactions

Example 5.4.13. *Consider a hypergraph with three vertices v_1, v_2, v_3, with a hyperedge h_1 with input v_1 and output v_1, v_2 and another hyperedge h_2 with input v_2, v_3 and output v_3 (Figure 5.4). Thus, v_1 and v_3 are catalysts. In this system, v_2 is created in h_1 with the help of v_1, without using up v_1, and it is destroyed in h_2 with the help of v_3, without creating anything. Each vertex appears once as input and once as output, and thus, this hypergraph represents a closed system of reactions in the sense of the definition. We shall call this a* source-sink system.

We shall now use this principle to create another example that is no longer a closed system of reactions, but makes use of the possibility demonstrated in the previous example to create and destroy products independently. And this will allow us to let the system branch and reunite in between.

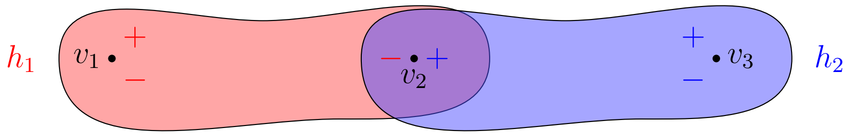


Figure 5.4: The hypergraph in Example 5.4.13.

Example 5.4.14. *Let Γ be the hypergraph with 4 vertices $v_1, \dots, v_4$ and 3 hyperedges h_1, h_2, h_3 such that (Figure 5.5):*

(i) h_1 has v_1 as input and v_2 as output;

(ii) h_2 has v_1 as output and v_3 as catalyst;

(iii) h_3 has v_1 as input, v_2 as input and v_4 as catalyst.

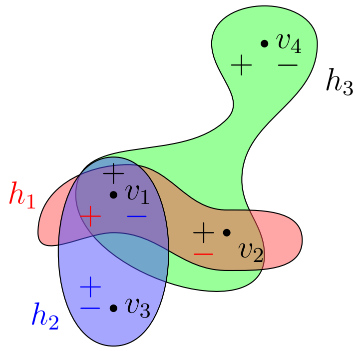


Figure 5.5: The hypergraph in Example 5.4.14.

This Γ does not contain any closed system. Now, let $\gamma : H \to \mathbb{R}$ such that $\gamma(h_1) := \gamma(h_3) := \frac{1}{2}$ and $\gamma(h_2) := 1$. Then γ satisfies (5.4.2), therefore $\lambda_1^H = 0$.

Proposition 5.4.15. *If Γ has l linearly independent closed systems, then*

$$\lambda_1^H = \ldots = \lambda_l^H = 0,$$

i.e. the multiplicity of the eigenvalue 0 for L^H is at least l.

Proof. Let $h_1 \dots, h_M$ be the hyperedges of Γ. If $\mathcal{S}_1, \dots, \mathcal{S}_l$ are linearly independent closed systems, it means that the rows of the $l \times M$ matrix $A = (a_{ij})_{ij}$ such that

$$a_{ij} := \begin{cases} 1 & \text{if } h_j \in \mathcal{S}_i \\ 0 & \text{otherwise} \end{cases}$$

are linearly independent. Therefore, the functions $\gamma_i : H \to \mathbb{R}$ defined as $\gamma_i(h_j) := a_{ij}$ for each $i \in \{1,\dots,l\}$ and each $j \in \{1,\dots,M\}$ are linearly independent. Also, they all satisfy (5.4.2). Therefore

$$\lambda_1^H = \ldots = \lambda_l^H = 0,$$

i.e. the multiplicity of the eigenvalue 0 for L^H is at least l. $\square$

Corollary 5.4.16. *If Γ has k pairwise disjoint closed systems, then*

$$\lambda_1^H = \ldots = \lambda_k^H = 0,$$

i.e. the multiplicity of the eigenvalue 0 *for L^H is at least k.*

Proof. The claim follows from Prop. 5.4.15 and from the fact that, if $\mathcal{S}_1,\dots,\mathcal{S}_k$ are pairwise disjoint closed systems, then they are also linearly independent. $\square$

5.4.1 Independent hyperedges and independent vertices

We end the section about the eigenvalue 0 by giving, with Prop. 5.4.20, another characterization of m_V and m_H. Before that, we define the *incidence matrix* of a hypergraph and we define *linear independence* for both hyperedges and vertices.

Definition 5.4.17. *Let $\Gamma = (V,H)$ be a hypergraph with N vertices $v_1,\dots,v_N$ and M hyperedges $h_1,\dots,h_M$. We define the $N \times M$ incidence matrix of Γ as $\mathcal{I} := (\mathcal{I}_{ij})_{ij}$, where*

$$\mathcal{I}_{ij} := \begin{cases} 1 & \text{if } v_i \text{ is an input and not an output of } h_j \\ -1 & \text{if } v_i \text{ is an output and not an input of } h_j \\ 0 & \text{otherwise.} \end{cases}$$

Therefore each row $\mathcal{I}_i$ of $\mathcal{I}$ represents a vertex v_i and each column $\mathcal{I}^j$ of $\mathcal{I}$ represents a hyperedge h_j.

Definition 5.4.18. *Given $J \subseteq \{1,\dots,M\}$, we say that the hyperedges $\{h_j\}_{j\in J}$ are linearly independent if the corresponding columns in the incidence matrix are linearly independent, that is, if $\{\mathcal{I}^j\}_{j\in J}$ are linearly independent. Analogously, given $I \subseteq \{1,\dots,N\}$, we say that the vertices $\{v_i\}_{i\in I}$ are linearly independent if the corresponding rows in the incidence matrix are linearly independent, that is, if $\{\mathcal{I}_i\}_{i\in I}$ are linearly independent.*

Remark 5.4.19. *Linear dependence of hyperedges means the following: we see each hyperedge as the sum of all its inputs minus the sum of all its outputs (and we can forget about the catalysts). If a hyperedge can be written as a linear combination of the other ones, with coefficients in $\mathbb{R}$, we talk about linear dependence. Analogously, in order to talk about linear dependence of vertices, we see each vertex as the sum of all the hyperedges in which it is an input minus the sum of all the hyperedges in which it is an output (and we can forget the hyperedges in which it is a catalyst).*
We also note that linear dependence does not depend on the choice of orientations for the hyperedges. When we change the orientation of the hyperedge h_j, the jth column of $\mathcal{I}$ is multiplied by -1, which does not affect linear dependence.

Proposition 5.4.20.

$$\dim(\ker \mathcal{I}) = m_H \qquad \text{and} \qquad \dim(\ker \mathcal{I}^\top) = m_V.$$

Proof. We first observe that we can see a function $\gamma : H \to \mathbb{R}$ as a vector $(\gamma_1, \dots, \gamma_M) \in \mathbb{R}^M$ such that $\gamma_j = \gamma(h_j)$. Also, two such functions are linearly independent if and only if the corresponding vectors are linearly independent. Now,

$$\begin{aligned}
\mathcal{I} \cdot \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_M \end{pmatrix} = \mathbf{0} &\iff \sum_{j=1}^{M} \mathcal{I}_{ij} \cdot \gamma_j = 0 \qquad \forall i \in \{1, \dots, N\} \\
&\iff \sum_{j_{\text{in}}: i \text{ input of } h_{j_{\text{in}}}} \gamma_{j_{\text{in}}} = \sum_{j_{\text{out}}: i \text{ output of } h_{j_{\text{out}}}} \gamma_{j_{\text{out}}} \qquad \forall i \in \{1, \dots, N\} \\
&\iff \gamma \text{ satisfies (5.4.2)} \\
&\iff \gamma \text{ is an eigenfunction of } L^H \text{ with eigenvalue } 0.
\end{aligned}$$

Therefore

$$\dim(\ker \mathcal{I}) = m_H.$$

With an analogous proof, one can see that

$$\dim(\ker \mathcal{I}^\top) = m_V.$$

□

We shall now see four corollaries of Prop. 5.4.20.

Corollary 5.4.21. *m_H and m_V don't change if we replace a hyperedge h containing a catalyst v by the new hyperedge $h \setminus \{v\}$.*

Proof. It follows from Prop. 5.4.20 and by the definition of $\mathcal{I}$. □

Corollary 5.4.22. *If there are linearly dependent hyperedges, then $m_H > 0$. If there are linearly dependent vertices, then $m_V > 0$.*

Corollary 5.4.23.

$$m_H = M - \textit{maximum number of linearly independent hyperedges}$$

and

$$m_V = N - \textit{maximum number of linearly independent vertices.}$$

Proof. It follows by Prop. 5.4.20 and by the Rank-Nullity Theorem. □

Corollary 5.4.24. *In the case of graphs,*

(i) Edges are linearly dependent if and only if they form at least one cycle;

(ii) Vertices are linearly dependent if and only if they cover at least one connected component of the graph.

Proof. In order to prove (i), assume first that a set of edges forms a cycle given by

$$e_1 = (v_1, v_2), e_2 = (v_2, v_3), \ldots, e_k = (v_k, v_1).$$

Then, if we consider the corresponding columns of the incidence matrix, it's clear that

$$\mathcal{I}^1 + \mathcal{I}^2 + \ldots + \mathcal{I}^k = 0.$$

Therefore any set of edges containing $e_1, \ldots, e_k$ is linearly dependent. Vice versa, assume that $e_1, \ldots, e_k$ are linearly dependent and let Γ' be the graph given by these edges. Then, by Corollary 5.4.22, $m'_H > 0$. Since m'_H is the number of cycles contained in Γ', this implies that $e_1, \ldots, e_k$ form at least one cycle.

One can prove (iii) in a similar way.

□

Remark 5.4.25. *Interestingly, the equation*

$$\mathcal{I} \cdot \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_M \end{pmatrix} = \mathbf{0} \tag{5.4.4}$$

for the eigenfunctions γ of L^H that have eigenvalue 0 *is analogous to the* metabolite balancing equation *in the* metabolic pathway analysis *[43]. In this setting, the v_i's are metabolites, the h_j's are metabolic reactions and the incidence matrix $\mathcal{I}$ is replaced by the similar* stoichiometric matrix*. With Equation (5.4.4), in this case, one looks for a* flux distribution $(\gamma_1, \ldots, \gamma_M)$ *such that each γ_j describes the net rate of the corresponding reaction h_j and such that, with this flux distribution, there is a balance between the metabolites which are consumed and the ones which are producted, in the overall stoichiometry. For this reason, Equation (5.4.4) in this case is called* metabolite balancing equation *and it describes the so-called* pseudo steady-state*. We should point out, however, that here we do not address the non-negativity conditions required in metabolic pathway analysis, as we are not working with directed hypergraphs.*

5.5 Largest eigenvalue

Since L^V and L^H are self-adjoint operators, we can apply the min-max principle and find, in particular, two alternative ways of computing λ_N:

(i)

$$\begin{aligned}
\lambda_N &= \max_f \frac{(\delta f, \delta f)_H}{(f,f)_V} \\
&= \max_f \frac{\sum_{h \in H} \delta f(h)^2}{\sum_{v \in V} \deg v \cdot f(v)^2} \\
&= \max_{f : \sum_{v \in V} \deg v \cdot f(v)^2 = 1} \sum_{h \in H} \delta f(h)^2 \\
&= \max_{f : \sum_{v \in V} \deg v \cdot f(v)^2 = 1} \sum_{h \in H} \Bigg(\sum_{v_i \text{ input of } h} f(v_i) - \sum_{v^j \text{ output of } h} f(v^j) \Bigg)^2
\end{aligned}$$

and

(ii)

$$\begin{aligned}
\lambda_N &= \max_{\gamma} \frac{(\delta^*\gamma, \delta^*\gamma)_V}{(\gamma,\gamma)_H} \\
&= \max_{\gamma} \frac{\sum_{v\in V} \deg v \cdot \delta^*\gamma(v)^2}{\sum_{h\in H}\gamma(h)^2} \\
&= \max_{\gamma:\sum_{h\in H}\gamma(h)^2=1} \sum_{v\in V} \deg v \cdot \delta^*\gamma(v)^2 \\
&= \max_{\gamma:\sum_{h\in H}\gamma(h)^2=1} \sum_{v\in V} \frac{1}{\deg v} \cdot \left(\sum_{h_{\text{in}}:v \text{ input}} \gamma(h_{\text{in}}) - \sum_{h_{\text{out}}:v \text{ output}} \gamma(h_{\text{out}}) \right)^2.
\end{aligned}$$

Example 5.5.1. *Consider a hypergraph with only one hyperedge h that involves N vertices: k inputs and m outputs, with $N \le k+m \le 2N$, so that there are $k+m-N$ catalysts. Then*

$$\begin{aligned}
\lambda_N &= \max_{\gamma:\gamma(h)^2=1} \sum_{v\in V} \left(\sum_{h:v \textit{ input}} \gamma(h) - \sum_{h:v \textit{ output}} \gamma(h) \right)^2 \\
&= \big|\textit{inputs that are not outputs}\big| + \big|\textit{outputs that are not inputs}\big| \\
&= \big|\textit{inputs}\big| + \big|\textit{outputs}\big| - 2\cdot\big|\textit{catalysts}\big| \\
&= k+m-2k-2m+2N \\
&= 2N-k-m.
\end{aligned}$$

In particular, this is the only eigenvalue of L^H. Observe that λ_N is equal to 0 if and only if $2N = k+m$, i.e. if and only if all vertices are catalysts, while λ_N achieves the largest value N exactly when $k+m=N$, i.e. when there are no catalysts.

Remark 5.5.2. *The previous example implies that λ_N cannot be bounded from above by a quantity that does not depend on the number of vertices N (while, for graphs, we always have $\lambda_N \le 2$). One should also compare this with Prop. 5.4.4.*

Bipartite hypergraphs

We know that the following theorem holds for graphs:

Theorem 5.5.3. *Let Γ be a connected graph. Then $\lambda_N \le 2$ and the equality holds if and only if Γ is bipartite.*

Recall 5.5.4. *Recall that a graph is bipartite if one can decompose the vertex set as a disjoint union $V = V_1 \sqcup V_2$ such that every edge has one of its endpoints in V_1 and the other in V_2.*

We will now generalize the notion of bipartite graph and extend it to hypergraphs, then we will generalize Theorem 5.5.3.

Definition 5.5.5. *We say that a hypergraph Γ is bipartite (Figure 5.6) if one can decompose the vertex set as a disjoint union $V = V_1 \sqcup V_2$ such that, for every hyperedge h of Γ, either h has all its inputs in V_1 and all its outputs in V_2, or vice versa.*

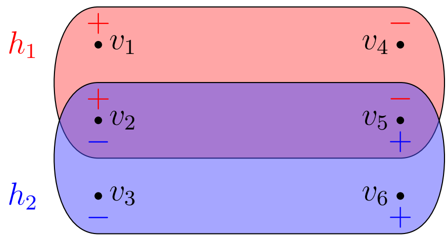


Figure 5.6: A bipartite hypergraph with $V_1 = \{v_1, v_2, v_3\}$ and $V_2 = \{v_4, v_5, v_6\}$.

Remark 5.5.6. *It is clear from the definition that:*

- *If a hypergraph is bipartite it does not contain catalysts;*
- *For a graph, our definition of bipartiteness reduces to the standard one of having two classes of vertices and no edges connecting vertices inside a class.*

Lemma 5.5.7. *Let Γ be a bipartite hypergraph. Then*

$$\lambda_N \geq \frac{\sum_{h \in H} |h|^2}{\sum_{h \in H} |h|}.$$

Proof. Since Γ is bipartite, we can write

$$\lambda_N = \max_{f:\sum_{v \in V} \deg v \cdot f(v)^2 = 1} \sum_{h \in H} \Bigg(\sum_{v_i \text{ input of } h} f(v_i) - \sum_{v^j \text{ output of } h} f(v^j) \Bigg)^2$$

$$= \max_{f:\sum_{v \in V} \deg v \cdot f(v)^2 = 1} \sum_{h \in H} \Bigg(\sum_{v_i \in h : f(v_i) > 0} f(v_i) - \sum_{v^j \in h : f(v^j) < 0} f(v^j) \Bigg)^2.$$

Now, let

$$f(v) := \frac{1}{\sqrt{\sum_v \deg v}}$$

for every $v \in V_1$ and

$$f(w) := -\frac{1}{\sqrt{\sum_w \deg w}}$$

for every $w \in V_2$. Then

$$\sum_{v \in V} \deg v \cdot f(v)^2 = 1$$

and

$$\sum_{h \in H} \Bigg(\sum_{v_i \in h : f(v_i) > 0} f(v_i) - \sum_{v^j \in h : f(v^j) < 0} f(v^j) \Bigg)^2$$

$$= \sum_{h \in H} \left(\frac{1}{\sqrt{\sum_v \deg v}} \cdot |h| \right)^2$$

$$= \frac{\sum_{h \in H} |h|^2}{\sum_{h \in H} |h|},$$

where the last equality is due to the fact that $\sum_v \deg v = \sum_{h \in H} |h|$. Therefore

$$\lambda_N \geq \frac{\sum_{h \in H} |h|^2}{\sum_{h \in H} |h|}.$$

□

Remark 5.5.8. *The quantity*

$$h' := \frac{\sum_{h \in H} |h|^2}{\sum_{h \in H} |h|}$$

appearing in Lemma 5.5.7 has the biggest value N exactly when every $h \in H$ has the biggest possible cardinality, which is N.

Remark 5.5.9. *Recall from Example 5.5.1 that, for bipartite hypergraphs with only one hyperedge, $\lambda_N = N$, therefore in this case $\lambda_N = h'$.*

Remark 5.5.10.

Apply Lemma 5.5.7 to a bipartite graph Γ. Since $|e| = 2$ for every edge, the lemma tells us that

$$\lambda_N \geq \frac{\sum_{e \in E} 4}{\sum_{e \in E} 2} = \frac{|E| \cdot 4}{|E| \cdot 2} = 2$$

and, as we know, this is actually an equality.

Proposition 5.5.11. *Let Γ be a hypergraph with largest eigenvalue λ_N. Then*

$$\lambda_N \leq \lambda'_N$$

where λ'_N is the largest eigenvalue of a bipartite hypergraph that has the same number of hyperedges as Γ and also the same number of inputs and the same number of outputs in each hyperedge (catalysts are not included).
The equality holds if and only if Γ is bipartite.

Proof. Let Γ be a hypergraph with largest eigenvalue λ_N. Then

$$\begin{aligned}\lambda_N &= \max_{f:\sum_{v\in V}\deg v\cdot f(v)^2=1} \sum_{h\in H}\Bigg(\sum_{v_i \text{ input of } h} f(v_i) - \sum_{v^j \text{ output of } h} f(v^j)\Bigg)^2 \\ &\leq \max_{f:\sum_{v\in V}\deg v\cdot f(v)^2=1} \sum_{h\in H}\Bigg(\sum_{v_i\in h: f(v_i)>0} f(v_i) - \sum_{v^j\in h: f(v^j)<0} f(v^j)\Bigg)^2,\end{aligned}$$

where the last inequality is due to the fact that, for every f,

$$\begin{aligned}&\Bigg|\sum_{v_i \text{ input of } h} f(v_i) - \sum_{v^j \text{ output of } h} f(v^j)\Bigg| \\ &\leq \Bigg|\sum_{v_i\in h: f(v_i)>0} f(v_i) - \sum_{v^j\in h: f(v^j)<0} f(v^j)\Bigg|.\end{aligned}$$

It is clear that the inequality for λ_N becomes an inequality if and only if, for every $h \in H$, we can let such f be positive in the inputs and negative in the outputs, or vice versa. And this is possible if and only if the hypergraph is bipartite. Therefore

$$\lambda_N \leq \lambda'_N$$

and the equality holds if and only if Γ is bipartite. □

Remark 5.5.12. *We can put together Lemma 5.5.7 and Prop. 5.5.11 and say that the largest value of λ_N is achieved by bipartite hypergraphs and that, in this case, $\lambda_N \geq h'$. In particular, $\lambda_N \geq h'$ becomes an equality for both bipartite graphs and bipartite hypergraphs with only one hyperedge. But it is in general not an equality, as proved by the next example.*

Example 5.5.13. *Let $\Gamma = (\{v_1, v_2, v_3, v_4\}, \{h_1, h_2\})$ be the bipartite hypergraph such that (Figure 5.7):*

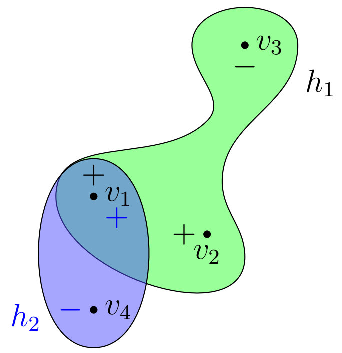


Figure 5.7: The hypergraph in Example 5.5.13.

(i) h_1 has v_1 and v_2 as inputs and v_3 as output;

(ii) h_2 has v_1 as input and v_4 as output.

In this case,

$$h' = \frac{\sum_{h\in H} |h|^2}{\sum_{h\in H} |h|} = \frac{13}{5} = 2,6.$$

Now we compute λ_N using the min-max principle applied to L^H. For simplicity, let $\gamma(h_1) := x$ and let $\gamma(h_2) := y$. Then

$$\begin{aligned}\lambda_N &= \max_{\gamma:\sum_{h\in H}\gamma(h)^2=1} \sum_{v\in V} \frac{1}{\deg v} \cdot \left(\sum_{h_{in}:v \text{ input}} \gamma(h_{in}) - \sum_{h_{out}:v \text{ output}} \gamma(h_{out}) \right)^2 \\ &= \max_{x,y\in\mathbb{R}:x^2+y^2=1} \left(x^2 + x^2 + \frac{(x+y)^2}{2} + y^2 \right) \\ &= \max_{x,y\in\mathbb{R}:x^2+y^2=1} \left(\frac{3}{2} + x^2 + xy \right),\end{aligned}$$

where in the last equality we have used the fact that $x^2 + y^2 = 1$. Now, let $x := \cos(t)$ and let $y := \sin(t)$. Then

$$\lambda_N = \max_{0\le t\le 2\pi} \left(\frac{3}{2} + \cos^2(t) + \cos(t)\cdot\sin(t) \right).$$

Now,

$$\frac{d}{dt}\left(\frac{3}{2} + \cos^2(t) + \cos(t)\cdot\sin(t) \right) = \cos(2t) - \sin(2t),$$

which has value 0 for $t = \frac{\pi}{8}$ and $t = \frac{5\pi}{8}$. In particular, for $t = \frac{5\pi}{8}$ we get that

$$\lambda_1^H = \frac{3}{2} + \cos^2\left(\frac{5\pi}{8}\right) + \cos\left(\frac{5\pi}{8}\right)\cdot\sin\left(\frac{5\pi}{8}\right)$$

$$= 2 - \frac{1}{\sqrt{2}}$$

$$\cong 1,29.$$

For $t = \frac{\pi}{8}$ we get that

$$\begin{aligned}\lambda_N &= \frac{3}{2} + \cos^2\left(\frac{\pi}{8}\right) + \cos\left(\frac{\pi}{8}\right) \cdot \sin\left(\frac{\pi}{8}\right) \\ &= 2 + \frac{1}{\sqrt{2}} \\ &\cong 2,71.\end{aligned}$$

In particular, $\lambda_N > h'$. This proves that the $\geq$ of Lemma 5.5.7 is, in general, not an equality.

We end this section by proving that there is another family of bipartite hypergraphs with $\lambda_N = h'$.

Lemma 5.5.14. *Let Γ be a bipartite graph on N nodes such that $\left|h\right| = N$ for every $h \in H$. Then*

$$\lambda_N = h' = N.$$

Proof. We first observe that, in this case,

$$h' = \frac{\sum_h \left|h\right|^2}{\sum_h \left|h\right|} = \frac{\left|H\right| \cdot N^2}{\left|H\right| \cdot N} = N.$$

Now observe that, for any bipartite hypergraph,

$$\begin{aligned}\lambda_N &= \max_f \frac{\sum_{h\in H}\left(\sum_{v_i \text{ input of } h} f(v_i) - \sum_{v^j \text{ output of } h} f(v^j)\right)^2}{\sum_{v\in V} \deg v \cdot f(v)^2} \\ &= \max_f \frac{\sum_{h\in H}\left(\sum_{v_i\in h: f(v_i)>0} f(v_i) - \sum_{v^j\in h: f(v^j)<0} f(v^j)\right)^2}{\sum_{v\in V} \deg v \cdot f(v)^2} \\ &= \max_{f>0} \frac{\sum_{h\in H}\left(\sum_{v\in h} f(v)\right)^2}{\sum_{v\in V} \deg v \cdot f(v)^2}.\end{aligned}$$

In our particular case, since $\{v \in h\} = \{v \in V\}$ for every h and since $\deg_v = |H|$ for every v, we have that

$$\begin{aligned}\lambda_N &= \max_{f>0} \frac{|H| \cdot \left(\sum_{v\in V} f(v)\right)^2}{|H| \cdot \sum_{v\in V} \cdot f(v)^2} \\ &= \max_{f>0} \frac{\left(\sum_{v\in V} f(v)\right)^2}{\sum_{v\in V} f(v)^2} \\ &= \lambda'_N,\end{aligned}$$

where λ'_N is the largest eigenvalue of a bipartite hypergraph on N nodes with only one hyperedge. As we have seen in Example 5.5.1, $\lambda'_N = N$, therefore $\lambda_N = h' = N$. □

5.6 Isospectral hypergraphs

We already know that two graphs cannot always be distinguished by their spectra, but the spectrum reveals some important properties. We expect something similar to happen for hypergraphs.

For instance, the spectrum of L of all complete bipartite graphs with the same number of vertices is the same [16]. (The multiplicity of the eigenvalue 0 of L^E, however, distinguishes between them.) For hypergraphs, a new phenomenon arises.

Lemma 5.6.1. *The spectrum of L^V and L^H doesn't change if we reverse the role of a vertex in all the hyperedges in which it is contained, i.e. if we let it become an input where it is an output and we let it become an output where it is an input.*

Proof. By the min-max principle, the spectrum of L^H is given by the *min-max* of the Rayleigh quotient, which is now

$$\frac{\sum_{v\in V} \frac{1}{\deg v} \cdot \left(\sum_{h_{\text{in}}:v \text{ input}} \gamma(h_{\text{in}}) - \sum_{h_{\text{out}}:v \text{ output}} \gamma(h_{\text{out}})\right)^2}{\sum_{h\in H} \gamma(h)^2}.$$

Now, since for each $v \in V$ we have

$$\left(\sum_{h_{\text{in}}:v \text{ input}} \gamma(h_{\text{in}}) - \sum_{h_{\text{out}}:v \text{ output}} \gamma(h_{\text{out}})\right)^2 = \left(\sum_{h_{\text{out}}:v \text{ output}} \gamma(h_{\text{out}}) - \sum_{h_{\text{in}}:v \text{ input}} \gamma(h_{\text{in}})\right)^2,$$

the Rayleigh quotient and therefore the spectrum of L^H (and L^V) doesn't change if we reverse the role of a vertex in all the hyperedges in which it is contained. $\square$

Example 5.6.2. *Let $\Gamma = (V, E)$ be a connected graph. Lemma 5.6.1 tells us that, if we reverse the role of a vertex $v \in V$ in all the edges in which it is contained, the spectrum of Γ doesn't change. This transformation actually creates an oriented graph where all edges that have v as an endpoint have either two inputs or two outputs. But this situation is not interesting from both the chemical point of view (where we assume that there are always both inputs and outputs) and the mathematical point of view, because in graph theory one always assigns an orientation to an edge by choosing exactly one input and exactly one output. Therefore, in order to have consistency with our theory, we should assume that every time we apply the operation described in Lemma 5.6.1 to a vertex v, we also apply it to all its neighbors. For the same reason, we should also apply it to the neighbors of its neighbors and therefore, by induction, since we are assuming that Γ is connected, we should apply this operation to all vertices of Γ. In conclusion, Lemma 5.6.1 in the case of graphs tells us that the spectrum doesn't change if we reverse the orientation of every edge in a given connected component.*

Remark 5.6.3. *Observe that the isospectral hypergraphs in Lemma 5.6.1 are not necessarily isomorphic in general. This is what makes the result interesting.*

6 Spectral measures and spectral classes

The content of this chapter is based on part of the results presented in the paper *Random geometric complexes and graphs on Riemannian manifolds in the thermodynamic limit*, a joint work with Antonio Lerario. Here we focus on the results regarding the Laplace operator for general graphs in order to remain coherent with the main topic of this thesis, and we refer the reader to [45] for a broader study of random geometric complexes and graphs.

6.1 Background

Fix a graph $\Gamma = (V, E)$ with spectrum

$$\lambda_1 = \lambda_1(\Gamma) \leq \ldots \leq \lambda_n = \lambda_n(\Gamma).$$

We define the *spectral measure* of Γ as

$$\mu_\Gamma := \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i(\Gamma)},$$

where δ denotes the Dirac measure. Since the eigenvalues are contained in the interval $[0, 2]$, μ_Γ is a probability measure on $[0, 2]$.
Jiao Gu, Jürgen Jost, Shiping Liu and Peter Stadler [27] introduced the notion of *spectral class* of a family of graphs, defined as follows. Given a Radon measure ρ on $[0, 2]$ and a sequence $(\Gamma_n)_{n\in\mathbb{N}}$ of graphs such that each Γ_n has n vertices, $(\Gamma_n)_{n\in\mathbb{N}}$ is said to belong to the spectral class ρ if

$$\mu_{\Gamma_n} \rightharpoonup \rho \text{ as } n \to \infty, \tag{6.1.1}$$

where the weak convergence in (6.1.1) means the following. A family of Radon measures μ_n on $[0, 2]$ converges weakly to the Radon measure μ_0, denoted

$$\mu_n \rightharpoonup \mu_0,$$

if for every continuous function $f : [0,2] \to \mathbb{R}$

$$\mu_n(f) \rightharpoonup \mu_0(f), \text{ as } n \to \infty.$$

In the particular case of $\mu_n = \mu_{\Gamma_n}$, we have

$$\mu_{\Gamma_n}(f) = \frac{1}{n}\sum_{i=1}^{n} f(\lambda_i(\Gamma)).$$

Quoting [27], the motivation for studying spectral measures and spectral classes comes from the fact that *empirical studies have shown that qualitatively different types of large graphs can in many cases be distinguished by the shape of their spectral density.* Summarizing the basic examples of spectral classes studied in [27],

- A graph on 2^n nodes is said to be a *cube graph* if its vertices and edges coincide with the vertices and edges of the n-dimensional unit cube [52]. Complete graphs, complete bipartite graphs and cube graphs asymptotically belong to the spectral class δ_1.
- Petal graphs asymptotically belong to the spectral class $\frac{1}{2}\delta_{1/2} + \frac{1}{2}\delta_{3/2}$.
- Paths and cycles asymptotically belong to the same spectral class ρ and this is such that $\rho(A) = 0$ for every finite subset $A \subset [0,2]$.
- Let $(\Gamma_n)_n$ be the family of graphs on n nodes defined as follows:

$$\Gamma_n = \begin{cases} K_n, & \text{if } n \text{ is even,} \\ \text{petal graph,} & \text{if } n \text{ is odd.} \end{cases}$$

  Then, there is no well-defined spectral class for $(\Gamma_n)_n$.

Also, the main result in [27] regarding spectral classes states the following. Define an *edit operation* on a graph as the insertion or deletion of an edge, or the insertion or deletion of an isolated vertex[1]. If two families of graphs differ by at most C edit operations *and* their corresponding spectral measures have weak limits, then they belong to the same spectral class. Formally,

[1]In this thesis, we assume that our graphs have no isolated vertices, i.e. vertices vertices of degree zero, because in this case $L = \mathrm{Id} - D^{-1}A$ and $\mathcal{L} = \mathrm{Id} - D^{-1/2}AD^{-1/2}$ are not well defined. However, in the literature, it is common to admit isolated vertices, with the additional convention that $(D^{-1})_{ii} = 0$ if $\deg v_i = 0$ [16].

Theorem 6.1.1 ([27], Theorem 2.8). *Let Γ_n and $\Gamma'_{n'}$ be graphs with n and n' vertices, respectively. Assume that $\Gamma'_{n'}$ can be obtained from Γ_n by at most C steps of edit operations, where C is independent of n. Then the families $(\Gamma_n)_n$ and $(\Gamma'_{n'})_{n'}$ belong to the same spectral class (assuming that the corresponding spectral measures possess weak limits).*

Here we prove an analogue of Theorem 6.1.1. Namely, in Theorem 6.2.3 we show that the difference of the spectral measure of two families of graphs differing by at most a finite number C of edges goes to zero weakly (without the assumption that the corresponding spectral measures have weak limits). We prove that our result holds not only for the Radon measures associated to L, but also for the ones associated to the adjacency matrix A, to the degree matrix D and to the non-normalized Laplacian matrix (or Kirchhoff matrix) K. Furthermore, in Section 6.2.2 we prove that, in the case of L, we have convergence in total variation distance for "connected sum" of complete graphs, but not for paths.

6.2 Main result

Definition 6.2.1. *Given an $n \times n$ matrix Q, denote the spectrum of Q by*

$$\lambda_1(Q) \leq \ldots \leq \lambda_n(Q).$$

We define the spectral measure of Q as

$$\mu_Q := \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i(Q)}.$$

Remark 6.2.2. *Observe that, since L and $\mathcal{L}$ have the same spectrum, for every graph Γ we have that*

$$\mu_{\mathcal{L}} = \mu_L = \mu_\Gamma.$$

Theorem 6.2.3. *Let $(\Gamma_{1,n})_n$ and $(\Gamma_{2,n})_n$ be two sequences of graphs such that, for every n, $(\Gamma_{1,n})_n$ and $(\Gamma_{2,n})_n$ are two graphs on n nodes that differ at most by C edges. Denote by $\mu_{1,n}$ and $\mu_{2,n}$ the spectral measures associated to one of the matrices A, D, K, $\mathcal{L}$. Then*

$$\mu_{1,n} - \mu_{2,n} \rightharpoonup 0,$$

i.e. for each continuous function $f : \mathbb{R} \to \mathbb{R}$

$$\left|\mu_{1,n}(f) - \mu_{2,n}(f)\right| \to 0.$$

We shall prove Theorem 6.2.3 in Section 6.2.1.

6.2.1 Proof of the main result

Preliminaries

Given a real $n \times n$ symmetric matrix Q, we define the 1-*Schatten norm of* Q as

$$\|Q\|_{S^1} := \sum_{i=1}^{n} |\lambda_i(Q)|.$$

The *Weilandt-Hoffman inequality* [57, Exercise 1.3.6] holds:

$$\sum_{i=1}^{n} |\lambda_i(Q_1) - \lambda_i(Q_2)| \leq \|Q_1 - Q_2\|_{S^1}. \tag{6.2.1}$$

We also define the *Frobenius norm of* Q [32, (5.6.0.2)] as

$$\|Q\|_F := \left(\sum_{i,j=1}^{n} |Q_{ij}|^2\right)^{1/2} = \sqrt{\operatorname{tr}(Q \cdot Q^\top)}.$$

Since Q is symmetric and since the trace of a matrix equals the sum of its eigenvalues [32, (1.2.4c)], we can write

$$\|Q\|_F = \sqrt{\operatorname{tr}(Q^2)} = \left(\sum_{i=1}^{n} \lambda_i(Q)^2\right)^{1/2}.$$

Proposition 6.2.4. *Let Q_1, Q_2 be real $n \times n$ symmetric matrices such that*

$$\|Q_1 - Q_2\|_{S^1} \leq C.$$

Then, for each $f : \mathbb{R} \to \mathbb{R}$ continuous and for each $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\left|\mu_{Q_1}(f) - \mu_{Q_2}(f)\right| \leq \varepsilon + \frac{2 \sup |f|}{\delta n} \cdot C.$$

Proof. Denote by $\{\lambda_i^{(1)}\}_{i=1}^n$ and $\{\lambda_i^{(2)}\}_{i=1}^n$ the eigenvalues of Q_1 and Q_2 respectively. Then

$$\mu_{Q_1}(f) - \mu_{Q_2}(f) = \frac{1}{n} \sum_{i=1}^{n} f(\lambda_i^{(1)}) - f(\lambda_i^{(2)}),$$

therefore

$$\left|\mu_{Q_1}(f) - \mu_{Q_2}(f)\right| \leq \frac{1}{n} \sum_{i=1}^{n} \left|f(\lambda_i^{(1)}) - f(\lambda_i^{(2)})\right|.$$

Now, since f is continuous, given $\varepsilon > 0$ there exists $\delta = \delta(f)$ such that

$$|\lambda_1 - \lambda_2| \leq \delta \qquad \Longrightarrow \qquad |f(\lambda_1) - f(\lambda_2)| \leq \varepsilon.$$

Therefore, since by (6.2.1) and by hypothesis we have that

$$\sum_{i=1}^{n} |\lambda_i^{(1)} - \lambda_i^{(2)}| \leq \|Q_1 - Q_2\|_{S^1} \leq C,$$

it follows that

$$\big|\{|\lambda_i^{(1)} - \lambda_i^{(2)}| > \delta\}\big| \leq \frac{C}{\delta}.$$

Therefore,

$$\begin{aligned}
&\Big|\mu_{Q_1}(f) - \mu_{Q_2}(f)\Big| \leq \frac{1}{n}\sum_{i=1}^{n}\big|f(\lambda_i^{(1)}) - f(\lambda_i^{(2)})\big| \\
&= \frac{1}{n}\sum_{|\lambda_i^{(1)}-\lambda_i^{(2)}|<\delta}\big|f(\lambda_i^{(1)}) - f(\lambda_i^{(2)})\big| + \frac{1}{n}\sum_{|\lambda_i^{(1)}-\lambda_i^{(2)}|\geq\delta}\big|f(\lambda_i^{(1)}) - f(\lambda_i^{(2)})\big| \\
&\leq \frac{1}{n}\sum_{|\lambda_i^{(1)}-\lambda_i^{(2)}|<\delta}\varepsilon + \frac{1}{n}\sum_{|\lambda_i^{(1)}-\lambda_i^{(2)}|\geq\delta}\big|f(\lambda_i^{(1)})\big| + \big|f(\lambda_i^{(2)})\big| \\
&\leq \frac{1}{n}\cdot\varepsilon\cdot\big|\{|\lambda_i^{(1)} - \lambda_i^{(2)}| < \delta\}\big| + \frac{1}{n}\cdot 2\sup|f|\cdot\big|\{|\lambda_i^{(1)} - \lambda_i^{(2)}| \geq \delta\}\big| \\
&\leq \frac{1}{n}\cdot\varepsilon\cdot n + \frac{1}{n}\cdot\frac{2\sup|f|}{\delta}\cdot C \\
&\leq \varepsilon + \frac{2\sup|f|}{\delta n}\cdot C.
\end{aligned}$$

□

Applications to graphs

Lemma 6.2.5. *Let Γ_1, Γ_2 be two graphs with $V(\Gamma_1) = V(\Gamma_2)$ that differ by at most C edges. Then,*

$$\|A_1 - A_2\|_{S^1} \leq 4C$$

$$\|D_1 - D_2\|_{S^1} \leq 4C^2$$

$$\|K_1 - K_2\|_{S^1} \leq 4C^2$$

$$\|\mathcal{L}_1 - \mathcal{L}_2\|_{S^1} \leq 2C\cdot\sqrt{2}\cdot\sqrt{n-1}.$$

Proof. Observe that any of the matrices

$$\Delta_1 := A_1 - A_2, \qquad \Delta_2 := D_1 - D_2, \qquad \Delta_3 := K_1 - K_2$$

consists of all zeros except for at most $4C$ entries, all of which entries are bounded by a constant (it is 1 for Δ_1 and C for Δ_2 and Δ_3). Therefore, each Δ_i, for $i = 1, 2, 3$ is an $n \times n$ real symmetric matrix that has rank at most $4C$. Since the rank of a square matrix equals the number of non-zero eigenvalues [47, Corollary 2.1.4], this implies that all eigenvalues of Δ_i are zero, except for at most $4C$ of them. It follows that, for $i = 1, 2, 3$, by first applying the Cauchy–Schwarz inequality,

$$\begin{aligned}
\|\Delta_i\|_{S^1} &= \sum_{j=n-4C+1}^{n} |\lambda_j(\Delta_i)| \\
&= \langle (|\lambda_{n-4C+1}|, \ldots, |\lambda_n|), (1, \ldots, 1) \rangle \\
&\leq \Big(\sum_{j=n-4C+1}^{n} \lambda_j(\Delta_i)^2 \Big)^{1/2} \sqrt{4C} \\
&= 2\sqrt{C} \|\Delta_i\|_F \\
&= 2\sqrt{C} \big(\sum_{j,k} (\Delta_i)_{jk}^2 \big)^{1/2} \\
&\leq 2\sqrt{C} (4C\tilde{C}_i^2)^{1/2} \\
&\leq 4C\tilde{C}_i,
\end{aligned}$$

where

$$\tilde{C}_i = \begin{cases} 1 & \text{if } i = 1 \\ C & \text{if } i = 2, 3. \end{cases}$$

Similarly, $\Delta_4 := \mathcal{L}_1 - \mathcal{L}_2$ consists of all zeros except for at most $2C(n-1)$ entries, all of which entries are bounded by 1, and it has rank at most $4C$. Therefore,

$$\begin{aligned}
\|\Delta_4\|_{S^1} &= \sum_{j=1}^{n} |\lambda_j(\Delta_4)| \\
&\leq \Big(\sum_{j=1}^{n} \lambda_j(\Delta_4)^2 \Big)^{1/2} \cdot \sqrt{4C} \\
&= \big(\sum_{j,k} (\Delta_4)_{jk}^2 \big)^{1/2} \cdot 2\sqrt{C} \\
&\leq \big(\sum_{1}^{2C(n-1)} 1 \big)^{1/2} \cdot 2\sqrt{C}
\end{aligned}$$

$$= 2C \cdot \sqrt{2} \cdot \sqrt{n-1}.$$

□

As a corollary, we can prove Theorem 6.2.3.

Proof of Theorem 6.2.3. We prove that, for each $f : \mathbb{R} \to \mathbb{R}$ continuous and for each $\varepsilon > 0$,

$$\lim_{n\to\infty} \left| \mu_{1,n}(f) - \mu_{2,n}(f) \right| \leq \varepsilon.$$

Let $c_1 := 4C$, $c_2 := c_3 := 4C^2$ and $c_4 := 2C \cdot \sqrt{2}$.
By Lemma 6.2.5, we have that

$$\|\Delta_i\|_{S^1} \leq c_i$$

for each $i = 1, 2, 3$. By Proposition 6.2.4, there exists $\delta > 0$ such that

$$\left| \mu_{1,n}(f) - \mu_{2,n}(f) \right| \leq \varepsilon + \frac{2\sup|f|}{\delta n} \cdot c_i.$$

Therefore,

$$\lim_{n\to\infty} \left| \mu_{1,n}(f) - \mu_{2,n}(f) \right| \leq \varepsilon.$$

Similarly, by Lemma 6.2.5 we have that

$$\|\Delta_4\|_{S^1} \leq c_4 \cdot \sqrt{n-1}.$$

By Proposition 6.2.4, there exists $\delta > 0$ such that

$$\left| \mu_{1,n}(f) - \mu_{2,n}(f) \right| \leq \varepsilon + \frac{2\sup|f|}{\delta n} \cdot c_4 \cdot \sqrt{n-1}.$$

Therefore,

$$\lim_{n\to\infty} \left| \mu_{1,n}(f) - \mu_{2,n}(f) \right| \leq \varepsilon$$

□

6.2.2 Strong convergence for complete graphs

In this section we prove that, in the case of L, we have convergence in total variation distance for "connected sums" of complete graphs, but not for paths. One could also ask what happens for the case of the Wasserstein distance, as studied in [28].

Definition 6.2.6. *We define the total variation distance between two probability measures μ_1 and μ_2 on $[0,2]$ as*

$$d_{\mathrm{tv}}(\mu_1,\mu_2) := \sup_{A\subseteq[0,2] \text{ measurable}} \Big|\mu_1(A)-\mu_2(A)\Big|.$$

Lemma 6.2.7. *Given $n\in\mathbb{N}_{>0}$, let K_n and K'_n be two complete graphs on n nodes. Let $K_n\sqcup K'_n$ be their disjoint union and let $\Gamma_n := K_n \bigcup_{C \text{ edges}} K'_n$ be their union together with C edges (v_i,v'_i) where $v_i\in K_n$ and $v'_i\in K'_n$, for $i=1,\dots,C$. Let also $\mu_{K_n\sqcup K'_n}$ and μ_{Γ_n} be the spectral measures of these two graphs. Then,*

$$\mu_{K_n\sqcup K'_n} = \frac{1}{n}\cdot\delta_0 + \left(\frac{n-1}{n}\right)\cdot\delta_{\frac{n}{n-1}}$$

and

$$\mu_{\Gamma_n} = \frac{1}{2n}\cdot\delta_0 + \frac{1}{2n}\cdot\sum_{i=1}^{2C+1}\delta_{a_i} + \left(\frac{n-1-C}{n}\right)\cdot\delta_{\frac{n}{n-1}}$$

for some $a_i\in(0,2)$.

Proof of Lemma 6.2.7. We use, in this proof, the notation L^{Γ} in order to indicate the Laplace operator for Γ.
Since the spectrum of $K_n\sqcup K'_n$ is given by 0 with multiplicity 2 and $\frac{n}{n-1}$ with multiplicity $2(n-1)$, we have that

$$\mu_{K_n\sqcup K'_n} = \frac{1}{2n}\left(2\cdot\delta_0 + 2(n-1)\cdot\delta_{\frac{n}{n-1}}\right).$$

In order to prove the second part of the lemma, we shall find $2(n-1-C)$ functions on $V(\Gamma_n)$ that are eigenfunctions for the Laplace operator with eigenvalue $\frac{n}{n-1}$ and are orthogonal to each other. In particular, by the symmetry of Γ_n, it suffices to find $n-1-C$ such functions that are 0 on the vertices of K'_n.

Observe that K_{n-C} is a subgraph of $K_n\setminus\{v_1,\dots,v_c\}$ that has $n-1-C$ eigenfunctions $f_1,\dots,f_{n-1-C}$ for the largest eigenvalue. These are orthogonal to each other and orthogonal to the constants, therefore

$$0 = \sum_{v\in V(K_{n-C})}\deg_{K_{n-C}}(v)f_i(v)f_j(v) = \sum_{v\in V(K_{n-C})} f_i(v)f_j(v)$$

and

$$\sum_{v\in V(K_{n-C})}\deg_{K_{n-C}}(v)f_i(v) = \sum_{v\in V(K_{n-C})} f_i(v) = 0$$

for each $i, j \in \{1, \ldots, n-1-C\}$. Now, for $i \in \{1, \ldots, n-1-C\}$, let $\tilde{f}_i$ be the function on $V(K_n)$ that is equal to zero on $v_1, \ldots, v_c$ and is equal to f_i otherwise. Then, $\tilde{f}_1, \ldots, \tilde{f}_{n-1-C}$ are orthogonal to each other and orthogonal to the constants because

$$\begin{aligned}\sum_{v \in V(K_n)} \deg_{K_n}(v) \tilde{f}_i(v) \tilde{f}_j(v) &= \sum_{v \in V(K_{n-C})} (n-1) \tilde{f}_i(v) \tilde{f}_j(v) \\ &= \sum_{v \in V(K_{n-C})} (n-1) f_i(v) f_j(v) \\ &= \sum_{v \in V(K_{n-C})} f_i(v) f_j(v) \\ &= 0\end{aligned}$$

and

$$\sum_{v \in V(K_n)} \tilde{f}_i(v) = \sum_{v \in V(K_{n-C})} f_i(v) = 0.$$

Since for complete graphs any function that is orthogonal to the constants is an eigenfunction for $\frac{n}{n-1}$, we have that $\tilde{f}_1, \ldots, \tilde{f}_{n-1-C}$ are (pairwise orthogonal) eigenfunctions for $\frac{n}{n-1}$.

Analogously, for $i \in \{1, \ldots, n-1-C\}$, let now $\hat{f}_i$ be the function on $V(\Gamma_n)$ that is equal to zero on $K_n' \cup \{v_1, \ldots, v_c\}$ and is equal to f_i otherwise. It's then easy to see that also these functions are orthogonal to each other and orthogonal to the constants. Now, for each i and for each $v \in \Gamma_n$ with $\tilde{f}_i(v) \neq 0$, we have that $v \in K_{n-C}$, therefore

$$\begin{aligned}L^{\Gamma_n} \hat{f}_i(v) &= \hat{f}_i(v) - \frac{1}{\deg_{\Gamma_n}(v)} \sum_{w \sim v} \hat{f}_i(w) \\ &= \tilde{f}_i(v) - \frac{1}{\deg_{K_n}(v)} \sum_{w \sim v \text{ in } K_n} \tilde{f}_i(w) \\ &= L^{K_n} \tilde{f}_i(v) \\ &= \frac{n}{n-1} \cdot \tilde{f}_i(v) \\ &= \frac{n}{n-1} \cdot \hat{f}_i(v).\end{aligned}$$

This proves that the functions $\hat{f}_i$ are $n-1-C$ orthogonal eigenfunctions of the Laplace Operator in Γ_n for the eigenvalue $\frac{n}{n-1}$. Since they are all 0 on K_n', by symmetry we can also get $n-1-C$ eigenfunctions for $\frac{n}{n-1}$ on Γ_n that are 0 on K_n and therefore are orthogonal to the first $n-1-C$ functions. This implies that the multiplicity of $\frac{n}{n-1}$ for Γ_n is at least $2(n-1-C)$. Therefore,

$$\mu_{\Gamma_n} = \frac{1}{2n} \cdot \delta_0 + \frac{1}{2n} \cdot \sum_{i=1}^{2C+1} \delta_{a_i} + \left(\frac{n-1-C}{n}\right) \cdot \delta_{\frac{n}{n-1}}$$

for some $a_i \in (0,2)$. □

Corollary 6.2.8. *The total variation distance between the probability measures $\mu_{K_n \sqcup K'_n}$ and μ_{Γ_n} defined in the previous lemma is*

$$\sup_{A \subseteq [0,2] \text{ measurable}} \left| \mu_{K_n \sqcup K'_n}(A) - \mu_{\Gamma_n}(A) \right| = \frac{2C+1}{2n}.$$

In particular, if $C = o(n)$, the total variation distance tends to zero for $n \to \infty$.

Example 6.2.9. *The previous corollary doesn't hold in general. As a counterexample, take two copies of the path on n vertices, P_n and P'_n. Their union via one external edge can be for example the path on $2n$ vertices, and*

$$\mu_{P_{2n}} = \frac{1}{2n} \left(\sum_{k=0}^{2n-1} \delta_{1-\cos\frac{\pi k}{2n-1}} \right),$$

while

$$\mu_{P_n \sqcup P'_n} = \frac{1}{n} \left(\sum_{k=0}^{n-1} \delta_{1-\cos\frac{\pi k}{n-1}} \right).$$

The total variation distance between these two measures does not tend to zero for $n \to \infty$.

Bibliography

[1] N. Alon and V. Milman. λ_1, isoperimetric inequalities for graphs, and superconcentrators. *J. Combin. Theory Ser. B*, 38(1):73–88, 1985.

[2] D. Balcan, H. Hu, B. Goncalves, P. Bajardi, C. Poletto, J. J. Ramasco, D. Paolotti, N. Perra, M. Tizzoni, W. Van den Broeck, V. Colizza, and A. Vespignani. Seasonal transmission potential and activity peaks of the new influenza A(H1N1): a Monte Carlo likelihood analysis based on human mobility. *BMC Medicine*, 7(45), 2009.

[3] A. Banerjee and J. Jost. On the spectrum of the normalized graph Laplacian. *Linear Algebra and its Applications*, 428:3015–3022, 2008.

[4] A. L. Barabási. *Linked: The New Science of Networks.* Perseus Books Group, 2002.

[5] A. L. Barabási. *Network Science.* Cambridge University Press, 2016.

[6] F. Bauer. Normalized graph Laplacians for directed graphs. *Linear Algebra and its Applications*, 436:4193–4222, 2012.

[7] F. Bauer, B. Hua, and J. Jost. The dual Cheeger constant and spectra of infinite graphs. *Advances in Mathematics*, 251:147–194, 2014.

[8] F. Bauer and J. Jost. Bipartite and neighborhood graphs and the spectrum of the normalized graph Laplacian. *Communications in Analysis and Geometry*, 21:787—-845, 2013.

[9] Y. Canzani. *Analysis on Manifolds via the Laplacian.* Notes, Harvard University, 2013.

[10] G. Carlsson, L. Dabaghian, Y.and Frank, and F. Mémoli. A topological paradigm for hippocampal spatial map formation using persistent homology. *PLoS Comput. Biol.*, 8(8), 2012.

[11] C. M. M. Casado. *La descrizione dell'universo con equazioni: Laplace*. RBA, 2017.

[12] K. C. Chang. The spectrum of the 1–Laplace operator. *Comm. Contemporary Mathematics*, 11:865–894, 2009.

[13] K. C. Chang. Spectrum of the 1–Laplacian and Cheeger's Constant on Graphs. *Journal of Graph Theory*, 81:167–207, 2016.

[14] G. Chartrand and P. Zhang. *A First Course in Graph Theory*. Dover Publications, 2012.

[15] J. Cheeger. A lower bound for the smallest eigenvalue of the Laplacian. *Princeton Univ. Press*, 1970.

[16] F. Chung. Spectral graph theory. *American Mathematical Society*, 1997.

[17] C. Curto, E. Gross, J. Jeffries, K. Morrison, M. Omar, Z. Rosen, A. Shiu, and N. Youngs. What makes a neural code convex? *SIAM J. Appl. Algebra Geom.*, 1(1), 2016.

[18] C. Curto and V. Itskov. Cell groups reveal structure of stimulus space. *PLoS Comput. Biol.*, 4(10), 2008.

[19] K. Das and S. Sun. Extremal graph on normalized Laplacian spectral radius and energy. *The Electronic Journal of Linear Algebra*, 29(1):237–253, 2016.

[20] J. Dodziuk. Difference equations, isoperimetric inequality and transience of certain random walks. *Trans. AMS*, 284, 1984.

[21] D. Easley and J. Kleinberg. *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge University Press, 2010.

[22] P. Erdős and A. Rényi. On random graphs. *Publicationes Mathematicae*, 6(290), 1959.

[23] E. Estrada. Graph and Network Theory in Physics. *arXiv:1302.4378*, 2013.

[24] L. Euler. Solutio problematis ad geometriam situs pertinentis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, 1736.

[25] C. Gordon, D. L. Webb, and S. Wolpert. One cannot hear the shape of a drum. *Bulletin of the American Mathematical Society*, 27(1), 1992.

[26] M. S. Granovetter. The strength of weak ties. *American Journal of Sociology*, 78(1360), 1973.

[27] J. Gu, J. Jost, S. Liu, and P. F. Stadler. Spectral classes of regular, random, and empirical graphs. *Linear Algebra Appl.*, 489:30–49, 2016.

[28] J. Gua, B. Hua, and S. Liu. Spectral distances on graphs. *Discrete Applied Mathematics*, 190-191:56–74, 2015.

[29] M. Hein and T. Bühler. An inverse power method for nonlinear eigenproblems with applications in 1–spectral clustering and sparse PCA. *NIPS*, pages 847–855, 2010.

[30] M. Hein and S. Setzer. Beyond spectral clustering – tight relaxations of balanced graph cuts. *Advances in Neural Information Processing Systems*, 24:2366–2374, 2011.

[31] D. Horak and J. Jost. Spectra of combinatorial Laplace operators on simplicial complexes. *Advances in Mathematics*, 244:303–336, 2013.

[32] R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, second edition, 2013.

[33] L. Isella, J. Stehlé, A. Barrat, C. Cattuto, J.-F. Pinton, and W. Van den Broeck. What's in a crowd? Analysis of face-to-face behavioral networks. *Journal of Theoretical Biology*, 271:166–180, 2011.

[34] D. Janežič, A. Miličević, S. Nikolić, and N. Trinajstić. *Graph Theoretical Matrices in Chemistry.* Mathematical Chemistry Monographs, 2007.

[35] J. Jost. *Riemannian geometry and geometric analysis.* Springer, seventh edition, 2017.

[36] J. Jost. *Mathematical Principles of Topological and Geometric Data Analysis.* In preparation, 2019.

[37] J. Jost and R. Mulas. Cheeger-like inequalities for the largest eigenvalue of the graph Laplace Operator. *arXiv:1910.12233*, 2019.

[38] J. Jost and R. Mulas. Hypergraph Laplace operators for chemical reaction networks. *Advances in Mathematics*, 351:870–896, 2019.

[39] J. Jost, R. Mulas, and F. Münch. Spectral gap of the largest eigenvalue of the normalized graph Laplacian. *Communications in Mathematics and Statistics*, 2019. Accepted.

[40] M. Kac. Can one hear the shape of a drum? *The American Mathematical Monthly*, 73(4):1–23, 1966.

[41] G. Kirchhoff. Ueber die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird. *Ann. Phys. Chem.*, 72, 1847.

[42] S. Kirkland. Completion of Laplacian integral graphs via edge addition. *Discrete Mathematics*, 295:75–90, 2005.

[43] S. Klamt and J. Stelling. Two approaches for metabolic pathway analysis? *Trends in Biotechnology*, 21, 2003.

[44] J. M. Lee. *Introduction to topological manifolds.* Springer, second edition, 2011.

[45] A. Lerario and R. Mulas. Random geometric complexes and graphs on Riemannian manifolds in the thermodynamic limit. *Discrete & Computational Geometry*, 2019. Accepted.

[46] J. Li, J.-M. Guo, and W. C. Shiu. Bounds on normalized Laplacian eigenvalues of graphs. *Journal of Inequalities and Applications*, 2014(1):316, 2014.

[47] T. Mach. *Eigenvalue Algorithms for Symmetric Hierarchical Matrices.* Doctoral dissertation, Chemnitz University of Technology, 2012.

[48] S. V. Maletić. *Simplicial complexes and complex networks: the influence of higher–order (sub)structures of network properties.* Doctoral dissertation, University of Belgrade, 2013.

[49] W. S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5:115–133, 1943.

[50] S. Minakshisundaram. Eigenfunctions on Riemannian manifolds. *J. Indian Math. Soc.*, 27:158–165, 1953.

[51] R. Mulas and N. M. Tran. Minimal embedding dimensions of connected neural codes. *Journal of Algebraic Statistics*, 2019. Accepted.

[52] S. Ovchinnikov. *Graphs and cubes.* Springer, 2011.

[53] A. F. Pascual. Advanced methods for dimensionality reductionand clustering: Laplacian eigenmaps and spectral clustering. Master's thesis, Universidad Autonoma de Madrid, 2010.

[54] G. Polya and S. Szegö. Isoperimetric inequalities in mathematical physics. *Annals of Math. Studies*, 27, 1951.

[55] O. Rojo and R. L. Soto. A new upper bound on the largest normalized Laplacian eigenvalue. *Operators and matrices*, 7:323–332, 2013.

[56] J.J. Sylvester. Chemistry and Algebra. *Nature*, 17, 1878.

[57] T. Tao. *Topics in random matrix theory*, volume 132 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2012.

[58] Oakland University. The Erdös Number Project. https://oakland.edu/enp/trivia/.

[59] H. Weyl. Über die asymptotische Verteilung der Eigenwerte. *Nachrichten der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, page 110–117, 1911.

[60] Wikipedia contributors. Page name: Simplicial complex. Wikipedia, The Free Encyclopedia.

[61] Wolfram Alpha LLC. WolframAlpha. https://www.wolframalpha.com.

Selbstständigkeitserklärung

Hiermit erkläre ich, die vorliegende Dissertation selbständig und ohne unzulässige fremde Hilfe angefertigt zu haben. Ich habe keine anderen als die angeführten Quellen und Hilfsmittel benutzt und sämtliche Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen wurden, und alle Angaben, die auf mündlichen Auskünften beruhen, als solche kenntlich gemacht. Ebenfalls sind alle von anderen Personen bereitgestellten Materialien oder erbrachten Dienstleistungen als solche gekennzeichnet.

Leipzig, den 11.12.2019

. . . .
(Raffaella Mulas)

Daten zum Autor

Name:	Raffaella Mulas
Geburtsdatum:	09.07.1992 in Cagliari (Italien)
09/2011 - 12/2014	Bachelorstudium in Mathematik Università degli Studi di Udine (Udine, Italien)
04/2015 - 07/2017	Masterstudium in Mathematik Rheinische Friedrich-Wilhelms-Universität Bonn (Bonn, Deutschland)
seit 09/2017	Doktorandin im Fachgebiet Mathematik Max-Planck Institut für Mathematik in den Naturwissenschaften, Universität Leipzig